MEDICINSKA FAKULTETEN
Lunds universitet
Institutionen för logopedi och foniatri

Underlying Factors in Speech Intelligibility: rationale, construction, and evaluation of a series of tests

Jonas Brännström

Audionomutbildningen, 2002
Vetenskapligt arbete, 20 poäng

Handledare: Dr Teresa Ching & Dr Harvey Dillon

CONTENTS

ABSTRACT 1

1. INTRODUCTION 2

1.1 Aim of study 2
1.2 Previous research on dead regions 2
1.3 Previous research on speech intelligibility 5
1.4 Proficiency and Speech Intelligibility Index (SII) 6
    1.4.1 Proficiency 7
    1.4.2 Speech Intelligibility Index (SII) 7
1.5 Summary 8

2. METHOD AND MATERIAL 9

2.1 Sessions and subjects 9
2.2 Equipment 9
2.3 Tests and stimuli 10
    2.3.1 Transiently Evoked Otoacoustic Emissions (TEOAE) 10
    2.3.2 Dead Regions' Test 10
    2.3.3 Psychophysical Tuning Curves (PTC) 11
    2.3.4 Filtered VCV-utterances and CUNY-sentences 12
2.4 Calibration 15
2.5 Data from the calibration 16
    2.5.1 First calibration measurement 16
    2.5.2 Second calibration measurement 19
    2.5.3 Derivation of correction figures 19
    2.5.4 Third calibration measurement 19
    2.5.5 Final calibration measurement 22
    2.5.6 Summary of calibration 24

3. RESULTS 25

3.1 Transiently Evoked Otoacoustic Emissions (TEOAE) 25
3.2 Dead Regions' Test 25
3.3 Psychophysical Tuning Curves (PTC) 30
3.4 CUNY-sentences and VCV-utterances 31
    3.4.1 CUNY-sentence test 31
    3.4.2 VCV's 33
    3.4.3 CUNY's and VCV's 35

4. DISCUSSION AND CONCLUSIONS 36

4.1 Transiently Evoked Otoacoustic Emissions (TEOAE) 36

4.2 Dead Regions' Test 36
4.3 Psychophysical Tuning Curves (PTC) 38
4.4 CUNY's and VCV's 38
    4.4.1 Noise (N) condition 38
    4.4.2 Quiet High (QH) condition 39
    4.4.3 Quiet Low (QL) condition 39
4.5 Summary 40
4.6 Acknowledgments 41

5. REFERENCES AND LITERATURE 42

5.1 References 42
5.2 Literature 43

ABSTRACT

The objective of the present study is to outline, generate, and evaluate a series of tests, which pertains to establish if there is a relationship between speech intelligibility and presence of dead regions. An extensive speech test was constructed using low-pass and band-pass filtered City University of New York- sentences (CUNY's) and vowel-consonat-vowel-utterances (VCV's). Furthermore, a test of frequency resolution, Psychophysical Tuning Curves (PTC), using simultaneous narrowband maskers with pulsing probe tones was generated. Together with two other tests, Dead regions' test (Moore et al., 2000) and Transiently Evoked Otoacoustic Emissions (TEOAE), they constitute the test setup. The stimuli and the equipment used were thoroughly evaluated and adjusted to minimize their impact on the results. Five normal hearing subjects participated in the present study That they indeed had normal hearing was established using screening with Pure Tone Audiometry (PTA). Four of the subjects participated in only the first part of the test, which included TEOAE, PTC, and the Dead regions' test. One subject participated in the first and the second part of the test. The second part included speech tests using CUNY's and VCV's in the different filtered conditions at two different levels in quiet and one in noise. The results indicated that the five subjects display normal outer hair cell function, normal PTA thresholds, and normal frequency resolution. The results from the Dead regions' test indicate that on average the Threshold Equalizing Noise test is working properly but there are individual differences. Generally, the results from the speech tests indicate that the scores correct increase with increasing cut-off frequency in the low-pass filtered conditions for both CUNY's and VCV's. In the same manner the speech scores increase with decreasing onset frequency in the band-pass filtered conditions. The conclusions were drawn that the series of tests are working as intended. However, in the speech tests an adjustment needs to be made regarding one of the two presentation levels in quiet to avoid obtaining scores correct that are too similar to the other.

1. INTRODUCTION

A common experience among audiologists in clinical practise is that two people with the same Pure Tone Audiogram (PTA) sometimes have such different speech identification abilities. The issue is that individual differences are most likely to impact on the optimal prescription if the people vary differently across frequency in their ability to extract information from audible signals.

The underlying hypothesis for this thesis is that speech intelligibility in an individual can be more effectively predicted from measurement of his/her frequency resolution and dead regions in the cochlea.

This work was conducted at the National Acoustic Laboratories (NAL) in Sydney, Australia. NAL is the research branch of Australian Hearing. The work at NAL lies in the audiological field with research in hearing assessment, hearing loss prevention, and rehabilitation devices and procedures.

1.1 Aim of study

Generally speaking this thesis accounts for the rationale and the construction of a series of tests aiming to obtain a better understanding of the underlying factors in speech intelligibility and hence gain accuracy in predicting speech intelligibility for hearing impaired people.

The objective of this study is selection of the tests used, construction of the tests, equipment setup, calibration, and evaluation of the results of a few normal hearing subjects.

This thesis tries to create a set of tests that provides us with data that can predict speech intelligibility from measurement of frequency resolution and dead regions in the cochlea. Furthermore this thesis aims to re-test the finding that proficiency is related to presence of dead regions (Vickers et al., 2001).

This study will mainly provide us with data from the construction of the test series and how it was done to make the tests run properly. This thesis is the first phase of a larger project, which will collect data from both hearing impaired and normal hearing subjects.

1.2 Previous research on dead regions

It has been suggested that differences in speech identification abilities may be due to the presence or absence of dead regions in the cochlea (Moore et al., 2000). The definition of a dead region is, according to Moore (2001),

> "when the IHC [inner hair cells] are non-functioning over a certain region of the cochlea, no transduction [via inner hair cells to neurones of the spiral ganglion] will occur in that region."

To pinpoint the exact location of a dead region in the cochlea he also states (Moore, 2001) that a dead region is defined

> "in terms of the CFs [characteristic frequencies] of the IHCs [inner hair cells] and/or neurones immediately adjacent to the dead region."

The latter is a more appropriate way to define a dead region, since it is still sufficient when the characteristic frequencies of the inner hair cells and neurones shift from their normal values, which Moore also mentions (Moore, 2001).

The presence of one or more dead regions in a cochlea results in what is been called *off-frequency listening*, which means that inner hair cells with other characteristic frequencies than what would normally correspond to in the input signal are used to detect it (Moore, 1995; O'Loughlin & Moore, 1981). This happens due to upward or downward spread of excitation depending on place of the dead region. When a complex sound, e.g. speech, enters the cochlea with a dead region, the inner hair cells of the region that is picking up the effect of the spread of excitation will have two incoming signals to transfer. This may effect the interpretation of the sound.

However, there is a problem with the presence of a dead region regarding the degree of "deadness" of a region. In the articles mentioned above it is assumed that a dead region is an all- or non-phenomenon, i.e. that the inner hair cells are completely dead in a dead region. It is likely that this assumption is a simplification. In a dead region there might be sections that are behaving as dead at low sensation levels due to severe damage, but at a certain level of sensation they might respond to an incoming signal. The response might be more or less distorted at these levels, but still information is transduced from the cochlea. The task of measuring the degree of "deadness" lies beyond the scope of this thesis, but the presumed implication of different degrees of "deadness" should be kept in mind while reading.

Dead regions are measured either by the use of Psychophysical Tuning Curves (PTC) or by pure-tones along with Threshold Equalising Noise (TEN) (Moore, 2001; Moore et al., 2001).

To measure Psychophysical Tuning Curves one uses a narrow-band noise with different centre frequencies as a masker for a pure tone at different frequencies (Pickles, 1995). The noise can be presented simultaneously with the stimulus or preceding the stimulus. The level of the pure tones is fixed at a chosen sensation level, preferably just above the threshold in quiet. The signal-to-noise ratio (SNR) required to mask the tone presented depends on whether the narrow-band noise is overlapping the pure tone stimulus or not (Moore, 2001; Pickles, 1995). If the noise is located far off in frequency the level to mask the pure-tone is going to be high, due to separation in frequency. If the noise is overlapping the tone presented then the level of the masker is going to be very low, since the noise and the pure-tone are overlapping (Fig. 1). This produces a *tip* in the plot at the stimulus frequency. In an inner ear with dead regions present the *tip* is shifted either upward or downward in frequency according to the upward or downward spread of excitation (Fig. 1). Which way the shift goes depends on the place of the dead region along the basilar membrane. This causes off-frequency listening. As mentioned above, an inner hair cell with a different characteristic frequency is used instead of the inner hair cell, which "normally" has the characteristic frequency of the incoming pure-tone frequency. Thus, the tip of the tuning curve lies at a different frequency that does not correspond to the signal input.

The Dead regions' test involves measurement of masked thresholds using a noise that is specified to give equally masked pure tone thresholds for persons with normal hearing (Moore et al., 2000). The noise is called Threshold Equalizing Noise (TEN). The noise level is specified as the level in a one-ERB (Equivalent Rectangular Bandwidth) (132Hz) wide band centred at 1000Hz (Moore, 2001; Vickers et al., 2001).

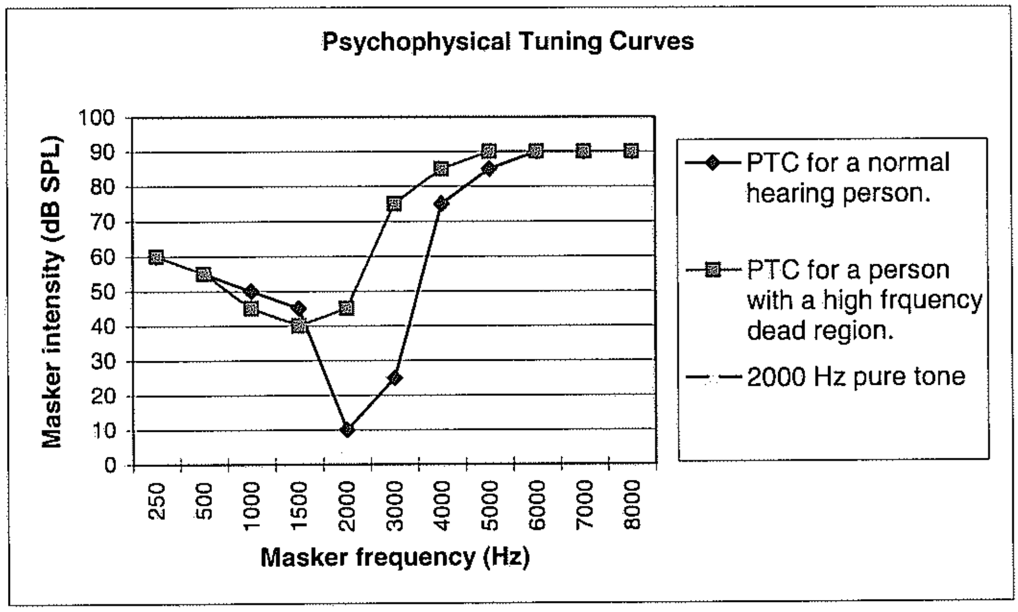

Figure 1. An illustration of how Psychophysical Tuning Curves (PTC) can look for one normal hearing person and one hearing-impaired person with a high frequency dead region. The masker is a narrow-band noise (1/3 critical band wide) with different centre frequencies. Stimulus is a 2000 Hz pure tone presented at 5 dB SL. The diamonds denote the PTC for the normal hearing person. The squares denote the PTC for the person with a high frequency dead region. The triangle denotes the pure tone stimulus.

The stimuli consist of pure tones. The Dead regions' test usually includes 11 points of measurement (250, 500, 1000, 1500, 2000, 3000, 4000, 5000, 6000, 8000 and 10000Hz). The noise presented is a broadband noise. The notion is that the broadband noise raises the threshold at the off-frequency listening point on the basilar membrane. This leads to a difference between the threshold in quiet and the masked one. When the masked threshold becomes 10 dB higher than the one measured in quiet and the masked threshold becomes 10 dB higher than the level of the masking noise, this indicates the presence of a dead region that corresponds with the place in frequency of the raised threshold (Moore et al., 2000). It has been shown that the shift of the PTC is highly correlated with this raise in threshold (Moore et al., 2000; Moore et al., 2001; Vickers et al., 2001).

1.3 Previous research on speech intelligibility

Ching et al. (1998) examined how amplification of speech affected listeners with different degrees of hearing loss. The averaged results showed that amplification of the high frequencies had less effect on the speech intelligibility with increasing hearing loss. This was achieved by comparing the speech scores of filtered sentences with predictions based on Speech Intelligibility Index (SII). Modifications were made to improve the predictability of SII, and they found that a method that combined the standard distortion factor with frequency-dependent proficiency factors derived for each person was more accurate in its predictions. The proficiency factors account for ability to extract information from an audible signal. However there were variations in the results of the new method that remained unexplained. The new SII formula underestimated the scores on intelligibility at low sensation levels for persons with severe to profound sloping hearing losses and overestimated the scores at high sensation levels. A throughout explanation of proficiency, the SII, and the calculations made by Ching et al. (1998) is to be found below in the section 1.4. Since results presented by Moore (2000) suggest that hearing losses greater than 85 dB nearly always are associated with dead regions, it seems likely to assume that the persons with severe and profound hearing losses could expose dead regions. This would perhaps account for the variations displayed.

Ching et al. (1997) examined the relationship between speech recognition, frequency resolution, temporal resolution, ability to use context, and age in hearing-impaired listeners. They found a decrease in frequency resolution and temporal resolution with increasing hearing loss, which might be explained by the high levels of intensity required by people with severe and profound hearing loss to make the signal audible. After they incorporated the standard level distortion factor to allow for the effects of absolute presentation level, they found that the individual proficiency factors was unrelated to frequency resolution (except at 4 kHz), temporal resolution, age and ability to use context. It seems likely to believe that the presence of dead regions could be correlated to both individual proficiency factors and frequency resolution. They also found that an audible signal could, for some people, affect the usefulness of a signal presented at a different frequency region so that it becomes useless or even decreases intelligibility. This is consistent with findings in persons with dead regions.

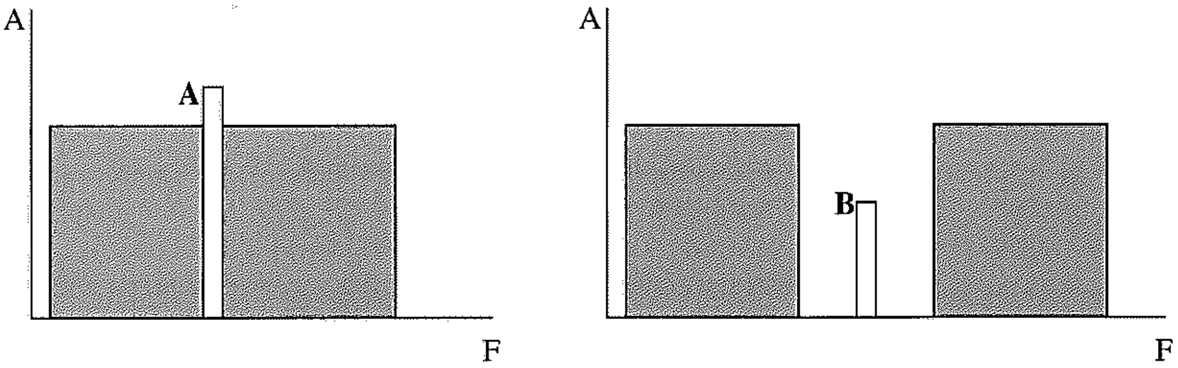

Figure 2. Illustration of the differences in threshold for a normal hearing subject, while using a broadband masker (left) and a notched masker (right). The difference in level between the two tones (**A** and **B**) is used to calculate the frequency resolution index. The shaded area depicts the noise applied. A stands for amplitude and F for frequency.

One way of calculating the ability of frequency resolution is by using the Frequency Resolution Index (FRI). The definition of FRI is the difference between the level of a narrowband masker and a notched-noise masker necessary to mask a probe tone (Ching et al., 1997), cf. Fig. 2. The level of the probe tone is fixed at a certain sensation level while the noise level is varied to find threshold. For example, B is the level required detecting a pure tone placed in a narrowband noise with a notch of 3 critical bands centred at the frequency of the pure tone stimulus. Let's say that the level B is 5 dB SL for a normal hearing subject. The level that is necessary to detect a pure tone in a narrowband noise without notch, we here call A. For now we assume that the level of A is 10 dB SL. This makes the Frequency Resolution Index in our hypothetical case, for one measured frequency, 5 dB. Thus, the formula for the calculation of the FRI is FRI = A-B. In the case of a hearing-impaired person without dead regions it is likely to believe that the value of the FRI will decrease towards zero with decreasing amount of hearing loss (Ching et al., 1997). However for a person with dead regions we expect to get FRI that is very close to zero or at zero due to off-frequency listening taking place outside the frequencies that are masked by the noise. Negative values of FRI were found in one study (Ching et al., 1997). This finding might be explained by the fact that the narrowband noise had narrower boundaries, i.e. cut-off frequencies, than the notched noise used (Ching et al., 1997), thus providing possibility for off-frequency listening in the narrowband noise condition but not in the notched noise condition.

A more accurate hearing aid prescription might be achieved taking dead regions into account, if proficiency is related to both dead regions and frequency resolution. We already know that dead regions are related to frequency resolution and the presence of dead regions to proficiency (Vickers et al., 2001). In their study (Vickers et al., 2001) they examined how speech perception for people with high frequency hearing losses with and without high frequency dead regions was affected when the speech stimuli, vowel-consonant-vowel (VCV) utterances, were low-pass filtered at different cut-off frequencies. The dead regions were diagnosed and the individual edges of the dead region were determined using PTC and Dead regions' test. The overall finding was that people without dead regions had continuously increasing speech scores with increasing frequency cut-off, while people diagnosed with dead regions only showed improvement in speech scores with increasing frequency cut-off until about 50-100% above the estimated edge of the dead region. Further increase of frequency cut-off lead to no improvement and in some cases decrease in the speech scores. The conclusions were drawn that people exhibiting high frequency hearing loss with dead regions do not benefit from amplification that exceeds the estimated edge frequency of the dead region with more than approximately 1.5-2 times.

The studies mentioned above were made in quiet and the results might not be applicable for e.g. speech intelligibility in noise.

1.4 Proficiency and Speech Intelligibility Index (SII)

A throughout explanation of the SII is made here although there are no actual calculations presented in this thesis. But since one of the aims of the overall project is to make these calculations, it was necessary to take them into account during the test construction.

1.4.1 Proficiency

Generally speaking, proficiency is a value of a person's ability to extract information from an audible signal depending on the enunciation of the speaker and the listener's experience with the speaker (Ching et al., 1998). This concept is based on the notion that the ability to extract information from an audible signal deteriorates with increasing hearing loss. For example, if the sensation level is raised, beyond a certain level, the person with the hearing loss will not benefit from it. The proficiency factor can be calculated in different ways depending on what we think is the better way to do it. The variables used can for example be ability of frequency resolution in different frequency regions compared with the ability to understand speech at different low-pass filtered cut-offs (e.g. Arlinger et al., 1990; Ching et al., 1998; Pavlovic et al., 1986). Proficiency is sometimes referred to as efficiency (e.g. Hogan & Turner, 1998).

1.4.2 Speech Intelligibility Index (SII)

As with its predecessor the Articulation Index (AI) (ANSI, 1969), the SII (ANSI, 1993) is based on the notion that speech intelligibility can be quantified. The SII quantifies the relationship between speech recognition and audibility and it assumes that when the audibility of a speech signal increases, the speech recognition also increases. The hearing thresholds of an individual are the foundation for this calculation along with the long term average spectra of the speech and noise that is reaching the ear of the individual. The original equation of the SII is as follows

$$\text{(eq. 1)} \qquad \mathrm{SII}=\Sigma\, I_i\, A_i.$$

The I stands for the importance of a certain frequency band, i, to speech intelligibility and the A indicates the part of the speech dynamic range that is above the listener's threshold (or masking noise) in a certain frequency band, i.

There have been made many evaluations of the SII and the general finding is that a decrease in speech recognition in hearing-impaired listeners (moderate, severe, or profound losses) can not be explained by audibility alone (e.g. Ching et al., 1998; Horwitz et al., 2001).

The ANSI 1993 Draft improved the correlation between the SII prediction and the actual values using a standard distortion factor. It calculated the standard level distortion factor to (as mentioned above) account for the decreasing benefit of audibility at high levels of stimulation in hearing-impaired people. Eq. 2 is the SII formula modified with the standard level distortion factor

$$\text{(eq. 2)} \qquad \mathrm{SII}=\Sigma\, I_i\, A_i L_i.$$

L_i is the standard level distortion factor, which is calculated in eq. 3

$$\text{(eq. 3)} \qquad L_i= 1-\frac{(E_i-U_i-10)}{160}.$$

E_i stands for the speech spectrum level and U_i for the standard speech spectrum for the normal vocal effort for the ith frequency band (ANSI S3.5 Draft v.3.1, 1993).

Ching et al. (1998) combined the SII improved by the standard level distortion factor with frequency-dependent proficiency factors derived for each person they tested. Eq. 4 expresses the modified formula:

$$\text{SII}=\Sigma\, I_i\, A_{I,s}\left(1-\frac{(E_i-U_i-10)}{160}\right)P_{n,s}, \qquad \text{(eq. 4)}$$

where the proficiency factor is denoted as $P_{n,s}$. More precisely this last expression is the proficiency factor P for the nth band for subject s. Using this new modified SII formula they could more accurately predict the effects of hearing loss on speech intelligibility. However, as mentioned earlier, the new formula underestimated the scores on intelligibility at low sensation levels and overestimated the scores at high sensation levels for persons with severe to profound sloping hearing losses (Ching et al., 1998). This discrepancy could not be accounted for by the new method and remained unexplained. This has a large impact on the prediction of speech intelligibility for the most common hearing loss, presbyacusis, which is characterized by a sloping envelope.

1.5 Summary

The results from the previous research mentioned indicates that dead regions are correlated with proficiency (Vickers et al., 2001). They also indicate that dead regions are correlated to one specific measurement of frequency resolution (Moore et al., 2001), i.e. PTC. Furthermore the results suggest a strong impact on speech intelligibility in an ear with dead regions (Moore, 2001; Vickers et al., 2001), but do not provide information about how we should take dead regions in to account when we try to predict speech intelligibility.

2. METHOD AND MATERIAL

The subjects, the equipment, the stimuli, and the procedures used are presented in this section. First, the generation of the stimuli and the procedures used in the tests are presented. After that the equipment used and the different calibration measurements are presented. Finally, the results from the calibration measurements and their implications are accounted for.

2.1 Sessions and subjects

This project was approved by the Australian Hearing Human Research Ethics Committee (HA02.1 - Factors affecting speech intelligibility in hearing-impaired people).

Five normal hearing subjects participated in the study. All subjects were native speakers of english and all younger than 30 years old. One subject participated in two sessions of two hours each. The other four participated in one session each.

Each session started with examination of the subject's ears with an otoscope to see that the meatus was not obstructed and the tympanic membrane looked normal.

The following test procedure was used. At the first session prior to any other testing transiently evoked otoacoustic emissions (TEOAE) were recorded. This was made to examine the function of the outer hair cells. After the otoacoustic emissions test Pure Tone Audiometry (PTA) was performed using the pure tones from the dead regions CD (Moore et al., 2000). PTA was measured to obtain thresholds in quiet. The PTA was followed by measurement of dead regions using the Dead regions' test and the PTC.

At the second session the speech intelligibility tests were performed using both City University of New York-sentences (CUNY) and vowel-consonant-vowel- utterances (VCV). On the third occasion was the same speech intelligibility tests performed but the order of the tests was reversed.

None of the subjects participated in all three sessions. This was due to lack of time. However, since the scope of this thesis is to account for rationale, setup, and evaluation of a series of tests, the results obtained were consistent for that purpose.

2.2 Equipment

Sennheiser HD 25-SP supra-aural headphones were used for the presentation of all the stimuli. The CD player used during the tests was a Yamaha CDX-530. Two equalizers were used of the model Behringer Ultra-Curve Pro 8024. Only one equalizer was used during the two first calibration measurements and the second one was added due to the results of these measurements. See section 2.5 for further explanation. Two attenuators were used in the test and they were Remote Control Attenuators. These attenuators were purpose built for National Acoustic Laboratories. A Marantz Integrated Stereo Amplifier PM-43 was used. A purpose built adder was also used.

The equipment was linked together in the following way (see Scheme 1). The two output channels of the CD were put through two separate channels of the first equalizer followed by the attenuators. The two separate outputs from the attenuators were fed into the adder. The adder joined together the two separate inputs into one channel output. The output from the adder passed through the Marantz to the earphones. The second equalizer was added before the first, thus between the CD player and the existing equalizer. This new equalizer was linked to the equipment also using two separate input and output channels.

This latter configuration was maintained and unaltered during the whole test. Once the equipment was linked together a set of calibrations measurements were made, see section 2.4 and 2.5 below.

2.3 Tests and stimuli

The conclusions drawn from the introduction in section 1.5 outlines the range of tests required. The following tests were used in this thesis; the PTA, the Dead regions' test, the Transiently Evoked Otoacoustic Emissions (TEOAE), the PTC, and the filtered speech tests.

2.3.1 Transient Evoked Otoacoustic Emissions (TEOAE)

TEOAE were measured using a GN Otometrics Capella. The presentation level of the stimuli used was 80 dB SPL (peak) which is the standard output level for this equipment. The non-linear mode was used. 1040 sweeps were sampled.

The parameters measured were waveform reproducibility (Waverepro %), emission strength (dB) and signal-noise ratio (dB) for frequency bands 500-1500 Hz, 1500-2500 Hz, 2500-3500 Hz, 3500-4500 Hz, and 4500-5500 Hz. The mean values for these three parameters were also registered.

NAL OAE software was used to calculate Coherent Emission Strength (CES). CES is a single number parameter calculated in dB SPL. It is the rms sound level of the emission, which is scaled according to the Waverepro in order to separate highly coherent records from highly "noisy" records. That is, it is a measure of the coherent part of the emission, or the "noise-free" part of the emission, which is common to both traces A and B of that record. It provides an output that is normally distributed for the NAL database population (LePage et al., 2001).

2.3.2 Dead regions' test

The PTA was measured using the pure tones from the Dead regions' test. The thresholds in quiet were obtained taking the average of 3-5 repetitions in an up and down fashion in 0.1 dB-steps using the purpose built attenuators. The subjects were instructed for both the pure tones in quiet and in noise to first adjust the level of the tone until they heard the tone clearly, then lower the level until they could not hear it anymore, adjust it until they just could hear it, and notify the test leader when they reached that level.

Dead regions were measured using the TEN. The noise level was, as mentioned earlier, specified as the level in a one-ERB (132 Hz) wide band centered at 1000 Hz. The pure tone stimuli were taken from the Diagnosis of Dead Regions CD (Moore, 2000). The TEN included 9 points of measurement (250, 500, 1000, 1500, 2000, 3000, 4000, 6000 and 8000 Hz). The method described by Moore et al. (2000) was used. The thresholds for this test were obtained using the same procedure as the one for the PTA.

It is worth noting that the pure tone stimuli on the Diagnosis of Dead Regions CD (Moore et al., 2000) are showing about 5 audible broadband frequency spikes each. This is probably due to how the stimuli on the CD were generated. But since one part of the aim of this project is to re-test the findings of Vickers et al. (2001) no changes were made to the pure tone stimuli.

2.3.3 Psychophysical Tuning Curves (PTC)

The measurement of frequency resolution was made using PTC, i.e. the level at which a pure tone at a fixed level is simultaneously masked by a narrowband noise.

The frequencies used for the probe tones were 500, 1000, 2000, and 4000 Hz. The pure tone stimuli were pulsing to facilitate the detection since the narrowband maskers somewhat resemble pure tones. The pure tone stimuli were generated in the software application Cool Edit 2000. The "purity" of the generated pure tones was verified with a 1/12-octave band analysis using a two-channel dynamic signal analyzer, Stanford Research Systems Model SR785. The stimuli were gated on for 300 msec and then followed by a 300 msec silent interval. The rise-fall time was 10 msec. The rms levels of the pure tone stimuli were equalized at –25 dBD prior to pulsing. 0 dBD equals digital saturation. This level is the same one used on the Speech and Noise CD for Hearing Aid Evaluation (2000) for the presentation of the stimuli.

The narrowband masker consisted of filtered white noise. The white noise used was generated by the two-channel dynamic signal analyzer, Stanford Research Systems Model SR785, and transferred into Cool Edit 2000. For each pure tone stimulus 7 center frequencies were selected - 0.24, 0.43, 0.78, 0.92, 1, 1.08, and 1.23 times the test frequency. The white noise was filtered with a bandwidth of 1/3 of the critical bandwidth (Zwicker & Terhardt, 1980) at these 7 center frequencies. However, the bandwidth was never allowed to be smaller than 80 Hz. This was made to avoid beats between the pure tone and the narrow-band masker. The extent of the filter slopes was kept constant at 10 Hz at each side of the band pass. Attenuation in the band stop was –70 dB and the measured slopes were all steeper than 50 dB per octave. The window function used for the filtering was Hamming. The narrow-band filters used are specified in Table 2.1. The filters were measured using 1/12 octave analysis and the same two-channel dynamic signal analyzer as mentioned above. The rms level of the noise was equalized at –25 dBD after passing through the filters.

In Cool Edit 2000 the pulsing pure tones and the filtered white noise were put onto one track but on two separate channels with the pulsing tone on the left channel and the noise on the right. Finally, the pure tone stimuli and the differently filtered white noise were put down on a CD as tracks of 2 minutes.

Table 1. Specification of the narrowband filters used for the measurement of Psychophysical Tuning Curves (PTC). The attenuation was set to -70 dB using 10 Hz wide slopes on each side of the band pass. CF stands for Center Frequency of the filter in Hz.

500 Hz probe tone		**1000 Hz probe tone**	
CF (Hz)	**Filter width (Hz)**	**CF (Hz)**	**Filter width (Hz)**
120	80	240	80
215	80	430	80
390	80	780	80
460	80	920	80
500	80	1000	80
540	80	1080	80
615	80	1230	80

2000 Hz probe tone		**4000 Hz probe tone**	
CF (Hz)	**Filter width (Hz)**	**CF (Hz)**	**Filter width (Hz)**
480	80	960	80
860	80	1720	80
1560	80	3120	150
1840	93	3680	183
2000	107	4000	233
2160	107	4320	233
2460	127	4920	300

The PTC session was made in the following way. The session started with a threshold measurement in quiet using the pulsing pure tones. The threshold was obtained by taking the average of 3-5 repetitions. Each pulsing pure tone stimulus (500, 1000, 2000, and 4000 Hz) was presented at 10 dB SL using the thresholds obtained in the quiet threshold measurement. The noise masker was introduced and the subjects adjusted the level of the masker until they just stopped to hear the pure tone stimulus. This procedure was repeated 3-5 times for each masker frequency depending on the consistency of the repetitions. i.e. if the subject did not manage to repeat the threshold without a difference exceeding two dB. The adjustments were made in 0.1-dB steps using the attenuators specified above.

2.3.4 Filtered VCV-utterances and CUNY-sentences

One part of the speech material consisted of 64 vowel-consonant-vowel (VCV) lists. The individual VCV's were taken from the Speech and Noise CD for Hearing Aid Evaluation (2000) and were randomly compiled into lists each consisting of 21 VCV's. Examples of VCV-utterances are *aga* and *ada*. The other part of speech material consisted of 70 City University of New York-sentences (CUNY) in an Australian version. An example of a CUNY-sentence is: *Have you eaten yet?*.

Filters were made using the Cool Edit 2000 and were specified as follows. Twenty seconds of pink noise was generated in Cool Edit 2000. The noise was measured using 1/12 octave analysis to confirm that it was indeed pink and had a flat spectrum.

The filters were constructed to have slopes steeper than 48 dB per octave, which was verified using a 1/12 octave analysis. The window function used for the filtering was Hamming. The equipment used for these analyses was the two-channel dynamic signal analyzer, Stanford Research Systems Model SR785, using 15 seconds of the generated pink noise. The stimuli, CUNY's and VCV's, were first, prior to any other filtering, filtered to match the International Long-Term Average Speech Spectrum (ILTASS) (Byrne et al., 1994) and then filtered to obtain seven speech bands: low-pass filtered at 700 Hz (LP 700), 1400 Hz (LP 1400), 2800 Hz (LP 2800), 5600 Hz (LP 5600); band-pass filtered at 690-5600 Hz (BP 690), 1380-5600 Hz (BP 1380), and 2760-5600 Hz (BP 2760). The slopes for each filter condition were achieved by using 1/10 of the onset and/or offset frequency, e.g. the slope for LP 700 was 70 Hz. The attenuation of the filter outside the slopes region was set to –70 dB. The noise consisted of the babble from 4 female and 4 male speakers filtered to match the ILTASS, which was taken from Speech and Noise CD for Hearing Aid Evaluation. However during the calibration measurement, it was discovered that the babble did exhibit slightly higher levels at some frequencies and it was decided that the levels of the speech stimuli should be adjusted accordingly after the babble noise. This last statement was later reconsidered (see section 2.5 below). The CUNY-sentences, the individual VCV, and the babble all were equalized at –25 dBD prior to any filtering. This level was, as mentioned earlier, the same as the one used on the Speech and Noise CD for Hearing Aid Evaluation (2000) and hence the level of the noise and the VCV-utterances used.

In Cool Edit 2000 the speech and the filtered babble were put onto one track for each list but on two separate channels with the speech on the left channel and the babble on the right.

The subjects began the first session involving speech tests practicing with running broadband speech (low-pass filtered at 5600Hz), 20-5600 Hz, presented in quiet. The running speech consisted of CUNY-sentences presented in a row with 0.5 seconds silent interval between them. It was decided that the shaping of the speech spectrum should be obtained using the POGO II formula, the prescription of gain and output, for the hearing impaired subjects to compensate audibility for the subjects' individual hearing loss. A high frequency emphasis was also introduced to the normal hearing subjects as well using the POGO II formula. A standard threshold was used for the calculation of the latter, which basically was -10 dB for 250 Hz, -5 dB for 500 Hz, and 0 dB for the following frequencies. The equalizer used for the shaping of the output was a Behringer Ultra-Curve Pro 8024.

Basically the original POGO formula is a modification of the half-gain rule, which takes the upward spread of masking in account and reduces the gain at 250 Hz and 500 Hz with –10 dB and –5 dB respectively. In the other frequencies this value is always 0 dB. The formula for POGO is as follows

$$IG_i = 0.5 * H_i + k_i$$

where IG_i stands for the insertion gain in a specified frequency, i, H_i for the hearing threshold at this specified frequency, and k_i for the constant that is added at the specified frequency as mentioned above (Dillon, 2001). The formula was later modified into POGO II, which is expressed as

$$IG_i = 0.5 * H_i + k_i$$

for hearing threshold less than 65 dB HL, and

$$IG_i = 0.5 * H_i + k_i + 0.5 * (H_i - 65)$$

for hearing losses greater than 65 dB HL. This was made since the original formula was intended for hearing losses up to 80 dB HL (Dillon, 2001). The modification simply adds 1 dB for each 1 dB increase in hearing loss above 65 dB HL. This means that the high frequency emphasis used for the normal hearing subjects was 0 dB for 250 Hz, 5 dB for 500 Hz, and 10 dB for the following frequencies used, 1000 Hz to 8000 Hz. The values for the 1/3-octave bands between 250 Hz and 500 Hz, and 500 Hz and 1000 Hz were interpolated.

The practice run with running speech was used to get the listeners familiarized with the speaker of the lists. The subject was also requested to adjust the level of the speech himself to get maximum intelligibility in the broadband LP 5600 by adjusting the attenuators to a level where the speech provided maximum intelligibility without being uncomfortably loud. After the level was set, a CUNY practice list was presented. Following this, the score correct was measured in these conditions with a new CUNY list. Another CUNY list was presented in the same broadband condition as above but now with the broadband babble noise added. The level of the noise was adaptively varied during the practice run to get 2/3 of the scores correct as obtained in the quiet condition described above. This signal-to-noise ratio was selected to avoid ceiling effects.

Thus, three different conditions were used, two in quiet and one in noise. The levels used for the quiet conditions were the levels of maximum clarity obtained from the results listening to the broadband speech (QH) and 1/3 of that level (QL). The latter level was not allowed to drop below 6 dB SL at first. However, this level was shown not to provide sufficient loudness for some of the filtered conditions used and a new lowest presentation level was adopted, 24 dB SL. This matter is accounted for in chapter 3, Results, and chapter 4, Discussion and conclusions. The level used for speech in noise (N) was the signal-to-noise ratio obtained during the practice run. These three levels for these three different conditions of level and noise were meant to be used at both speech sessions.

Table 2. Specification of the filters used for the filtering the CUNY-lists and the VCV-lists.

Filter	Offset attenuation -70 dBD	Onset band-pass 0 dBD	Offset band-pass 0 dBD	Onset attenuation -70 dBD
LP 700 Hz	-	0 Hz	700 Hz	770 Hz
LP 1400 Hz	-	0 Hz	1400 Hz	1540 Hz
LP 2800 Hz	-	0 Hz	2800 Hz	3080 Hz
LP 5600 Hz	-	0 Hz	5600 Hz	6160 Hz
BP 690 Hz	621 Hz	690 Hz	5600 Hz	6160 Hz
BP 1380 Hz	1242 Hz	1380 Hz	5600 Hz	6160 Hz
BP 2760 Hz	2484 Hz	2761 Hz	5600 Hz	6160 Hz

Speech scores were measured for the different filter types (specified above in table 2) and for the three different conditions, QH, QL, and N. The order of the filters and the level or noise conditions was counterbalanced between subjects and the two sessions using the Latin square and the inverse Latin square. Practice using running speech was made for each filter type. After the subjects had practiced with running speech, a practice list of CUNY-sentences was presented. The "real" test list was then presented and scores correct were obtained for this particular condition and filter type. The testing was repeated for each level or noise condition in the same filter type. This whole procedure was repeated for all the three conditions mentioned above and in the same manner for the VCV's. This means that the CUNY's and the VCV's were tested in a row using the same filter type. After all the conditions were tested for one filter type, a new filter type was tested following the same procedure. This was repeated for all filter types. An exact replication of these tests was meant to be made during the second session although the conditions and filter types were presented in the opposite order to accomplish counterbalancing.

It was intended that all subjects were to be tested in the same manner as described above. Only the order of the different level or noise conditions and filter types was to be altered. Since only one subject was tested at one session only, the counterbalancing has no effect on the results presented in this thesis, but it will in the larger project mentioned.

2.4 Calibration

The calibration was made to determine how the equipment worked and the absolute levels in dB SPL of the different stimuli passing through the linked equipment. The results from the calibration measurements are presented in section 2.5.

The earphones were calibrated in a 6-cc coupler with a Brüel and Kjaer 4144 pressure microphone using a Brüel and Kjaer Frequency Analyser Type 2121 (ISO 389, 1998), which itself was calibrated to a Brüel and Kjaer Pistonphone Type 4220 at 124 dB SPL. The weight used to add a nominal static pressure to the earphones was 455g (ISO 389, 1998). The calibration of the attenuators was checked with the Brüel and Kjaer Frequency Analyser Type 2121. Furthermore, KEMAR was used along with a Brüel and Kjaer 4192 pressure microphone. The calibration took place in a soundproof both.

The calibration was performed as follows. The master gain of the equalizer was set to +12 dB at both channels, left and right. The attenuation of the attenuators was set to 20 dB on one, which means that it lowered the level with 20 dB, and 99.9 on the other. A calibration tone, 1000 Hz, taken from the Speech and Noise CD for Hearing Aid Evaluation (2000) was presented through the earphones and the level was measured with the Brüel and Kjaer Frequency Analyzer and the attachments mentioned above. The same procedure was repeated but now with the calibration noise, also taken from Speech and Noise CD for Hearing Aid Evaluation (2000). The 1/3 octave spectrum of the calibration noise were measured using the Brüel and Kjaer Frequency Analyzer connected to a two-channel dynamic signal analyzer, Stanford Research Systems Model SR785. The 1/3 octave analysis was made both through the earphones and direct from the CD player. This was done to obtain the difference between the original signal and the one passed through the equipment. The 1/3 octave spectra were also measured for the different stimuli used to obtain the absolute values coming out of the linked equipment and earphones. This means that the 1/3 octave spectra of the noise and the pure tone stimuli of the Dead regions' test, the narrowband maskers and the pulsing pure tone stimuli of the PTC, and the speech maskers and the running speech for the filtered conditions used, were measured.

To obtain the absolute level in dB SPL the level of each 1/3 octave band, the values in dBV rms indicated by the SR785 spectrum analyzer were added to the reading from the display of the Brüel and Kjaer Frequency Analyzer and from that sum 10 was subtracted. The reason for this correction was that when the input to the Brüel and Kjaer Frequency Analyser equaled its displayed setting, the voltage out of the Brüel and Kjaer Frequency Analyser (in dBV) was 10 dB greater than the displayed setting (in dB SPL). Each stimulus was subjected to the 1/3 octave analysis using a 30 seconds timeframe. The frequency range extended from 50 Hz to 10000 Hz. The linearity of the equipment was also examined by measuring the output of the earphones at different levels on the attenuators and the amplifier.

2.5 Data from the calibration

Here the results from the calibration are presented and the alterations made are accounted for.

2.5.1 First calibration measurement

The results from the first calibration measurement using pink noise are presented in Table 3 and Figure 3. Four levels were established on the amplifier using the 1000 Hz calibration tone. The first step was adding 0 dB, the second 20 dB, the third 40 dB, and the fourth 45 dB. These different amplifier settings were obtained to extend the range of the equipment. The output from the CD player was measured directly into the spectrum analyzer resulted in a fairly flat response, which is expected using a 1/3 octave analysis. However, when the pink noise was passed through the equipment and the earphones, the result was a non-flat response with higher levels in the frequency bands 50-250 Hz ranging from 4.2 dB to 11.2 dB, 1600-3150 Hz ranging from 3.4 dB to 5.2 dB, and 6300-10000 Hz ranging from 2.3 dB to 6.1 dB. The levels were lower in the frequency bands 400-630 Hz, on average about 1.5 dB. There was also a dip, -2.3 dB, in the 5000 Hz frequency band. The increased levels in the lower frequency bands 50 Hz to 250 Hz were probably due to the fact that a 6 cc-coupler was used with a weight, which rather is adapted to TDH 39 earphones than to the Sennheiser HD 25-SP supra-aural headphones used in this thesis.

Table 3. 1/3 octave analysis of pink noise, prior to any adjustments, through earphones using 6-cc coupler and CD direct into the spectrum analyzer, Stanford Research Systems Model SR785. The right attenuator was set to 99.9, the left to 20.0, and the amplifier at level 3. Differences >1 dB are marked in italic.

1/3 octave bands (Hz)	Pink noise through earphones (dB SPL)	Pink noise CD direct, eq at 1000 Hz (dB SPL)	Difference (dB)
50	67,7	66,6	*1,1*
63	75	67,6	*7,4*
80	77,4	66,9	*10,5*
100	78,5	67,3	*11,2*
125	77,9	67,4	*10,5*
160	75,8	66,8	*9*
200	74,3	67,3	*7*
250	71,3	67,1	*4,2*
315	67,6	66,9	0,7
400	65,8	67,3	*-1,5*
500	65,5	67,1	*-1,6*
630	65,6	67,1	*-1,5*
800	66,2	67,2	-1
1000	67,3	67,3	0
1250	67,9	66,9	1
1600	70,4	67	*3,4*
2000	71,6	67,1	*4,5*
2500	71,8	66,6	*5,2*
3150	71,4	66,8	*4,6*
4000	68,1	67	*1,1*
5000	64,2	66,5	*-2,3*
6300	70,1	66,6	*3,5*
8000	72,8	66,7	*6,1*
10000	68,5	66,2	*2,3*

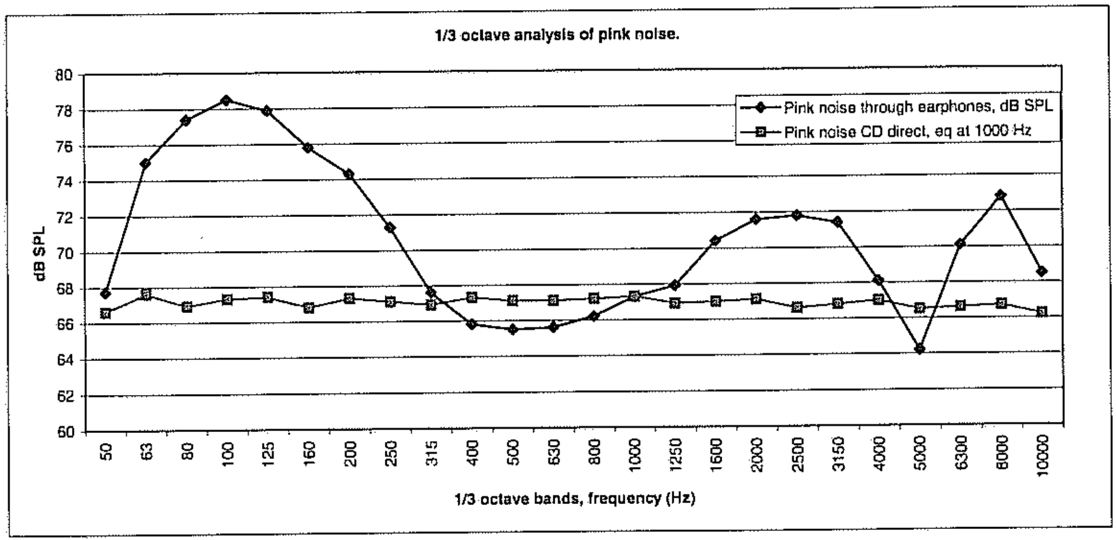

Figure 3. 1/3 octave analysis of pink noise prior to any adjustments. The squares denote the pink noise measured direct from the output of the CD player. The diamonds denote the pink noise measured through all the equipment and the earphones using a 6-cc coupler and a weight of 455 grams. The right attenuator was set to 99.9, the left to 20.0, and the amplifier at level 3. The levels are equalized at 1000 Hz.

The boost in the frequency bands 1600-3150 Hz and the dip at 5000 Hz was a real

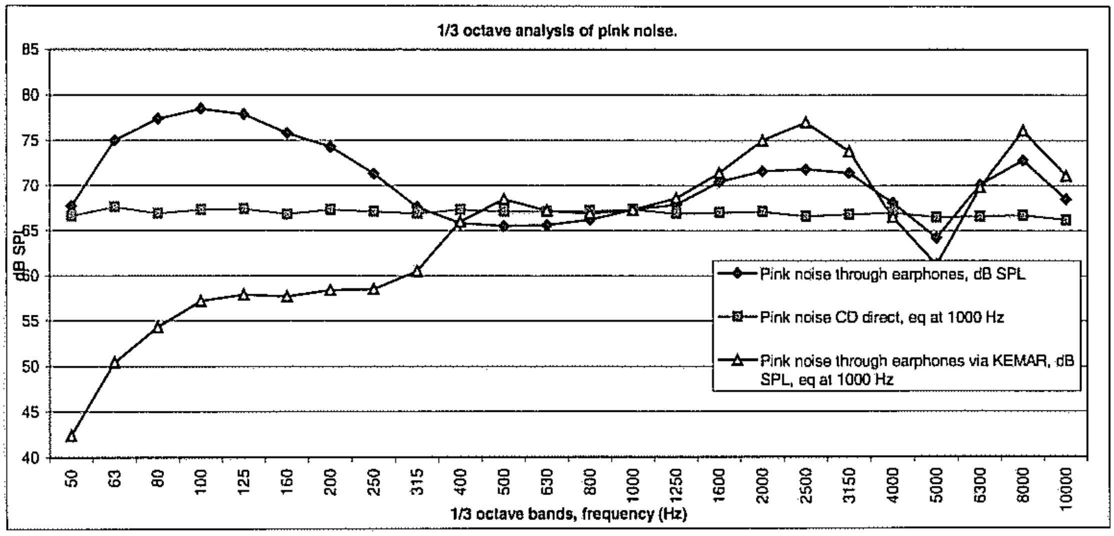

Figure 4. 1/3 octave analysis of pink noise prior to any adjustments. The squares denote the pink noise measured direct from the output of the CD player. The diamonds denote the pink noise measured through all the equipment and the earphones using a 6-cc coupler and a weight of 455 grams. The triangles denote the pink noise measured through all the equipment and the earphones using KEMAR. The right attenuator was set to 99.9, the left to 20.0, and the amplifier at level 3. The levels are equalized at 1000 Hz

Table 4. 1/3 octave analysis of pink noise, prior to any adjustments, through earphones using KEMAR and CD direct into the spectrum analyzer, Stanford Research Systems Model SR785. The right attenuator was set to 99.9, the left to 20.0, and the amplifier at level 3. Differences >1 dB are marked in italics.

1/3 octave bands (Hz)	Pink noise through earphones via KEMAR (dB SPL)	Pink noise CD direct, eq at 1000 Hz (dB SPL)	Difference (dB)
50	42,4	66,6	*-24,2*
63	50,4	67,6	*-17,2*
80	54,3	66,9	*-12,6*
100	57,2	67,3	*-10,1*
125	57,9	67,4	*-9,5*
160	57,7	66,8	*-9,1*
200	58,4	67,3	*-8,9*
250	58,5	67,1	*-8,6*
315	60,5	66,9	*-6,4*
400	66	67,3	*-1,3*
500	68,5	67,1	*1,4*
630	67,2	67,1	0,1
800	66,9	67,2	-0,3
1000	67,3	67,3	0
1250	68,6	66,9	*1,7*
1600	71,4	67	*4,4*
2000	75	67,1	*7,9*
2500	77	66,6	*10,4*
3150	73,8	66,8	*7*
4000	66,5	67	-0,5
5000	61,2	66,5	*-5,3*
6300	69,8	66,6	*3,2*
8000	76,1	66,7	*9,4*
10000	71,1	66,2	4,9

problem since the cut-off frequency for four of the filtered conditions was 5600 Hz. The difference between 1/3 octave bands 2500 Hz and 5000 Hz was 7 dB. This meant that, for example, the filtered condition BP 2800 would basically not match the ILTASS and hence give a lower contribution to speech intelligibility than expected. Since the 6-cc coupler starts to resonate somewhere above 5000 Hz, it was decided that another measurement of the pink noise should be done using the real ear simulator KEMAR to determine if the resonance vanished.

2.5.2 Second calibration measurement

In Table 4 and Figure 4 are the results from the second calibration measurement presented. Instead of getting a low frequency boost as in the previous measurement using the 6-cc coupler, we now got a lowering in the levels of the lower frequency bands, 50-315 Hz, using KEMAR. It seems as if the discussion about the influence of the weight used proved right, since the amount of pressure applied by the headband of the Sennheiser HD 25-SP is lower than the one of the TDH 39. Furthermore, an accentuation of the high levels in the frequency bands 1600-3150 Hz and 6300-10000 Hz is shown. The dip at 5000 Hz has also become deeper and we now have a difference of 15.1 dB between the frequency bands 2500 Hz and 5000 Hz. This was aggravating the problem mentioned in the previous section.

2.5.3 Derivation of correction figures

To flatten out the response of the earphones it was decided that the equalizer settings should include corrections for this discrepancy from the CD's direct pink noise. The values were derived from the comparison between the measurement using the 6-cc coupler and the CD's direct noise. The values were applied to the equalizer. The effects of the correction values in the equalizer are presented in the following section below. The differences between the levels of the pink noise in for example Figures 4 and 5 are only due to the fact that the levels are equalized at the frequency band 1000 Hz using KEMAR in the previous and the 6-cc coupler in the latter as reference.

The argument for taking the correction values from the measurement using the 6-cc coupler was that there is no point in trying to flatten out the natural resonance of the outer ear and the meatus, which would more or less have been the case if the values from the KEMAR measurement would have been used. It was also decided that another equalizer should be added. The amount of corrections in the equalizer and the overall level, +12 dB, was decreasing the capability to provide the spectrum shaping needed for the subjects with sloping hearing loss.

2.5.4 Third calibration measurement

From this point the second equalizer was linked to the equipment in the manner described earlier. The levels in this equalizer were initially set to zero flat in all frequencies. It was thus not bypassed. The chain of equipment was measured and the results, not presented in this study, clearly showed that the new equalizer did not in any respect alter the output of the earphones.

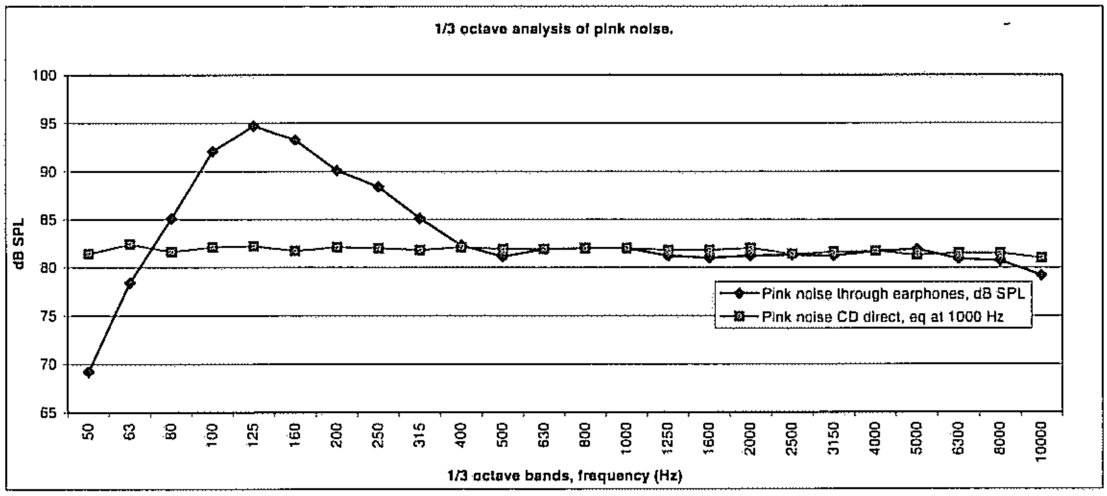

Figure 5. 1/3 octave analysis of pink noise after the correction for the deviant levels was made. The right attenuator was set to 99.9, the left to 20.0, and the amplifier at level 3. The squares denote the pink noise measured direct from the output of the CD player. The diamonds denote the pink noise measured through all the equipment and the earphones using a 6-cc coupler and a weight of 455 grams. The levels are equalized at 1000 Hz.

The results from the third calibration measurement using the 6-cc coupler again are presented below in Figure 5. In the frequency bands 400-8000 Hz there were no deviations from the pink noise lined directly into the spectrum analyzer larger than 0.8 dB. There were still higher levels in the lower frequencies 80-315 Hz, except for 50 Hz and 63 Hz, which now had dropped below the CD direct input. The difference in the results for the lower frequency bands seems, as mentioned above, to be depending on whether the 6-cc coupler or the KEMAR is used. It was decided that the levels in these low frequencies really do not matter as long as they are not providing any upward spread of masking due to high levels. That does not seem to be the case. The comparison between the results from the measurement using KEMAR and the 6-cc coupler indicate that these levels that are fairly high in the 6-cc coupler will drop when the earphones are fitted onto a subjects ears. Furthermore, the only speech signal contained in this frequency region is - more or less depending on if it is a male or female speaker - the fundamental frequency, but all information pertaining from it can be derived from its formants. That is why these high levels in the 6-cc coupler were tolerated, although this might impose some effects on the results when measuring the 250 Hz pure tone in the dead regions test regarding the absolute level of the threshold but not in comparison with the other subjects in the study.

The frequency band 10000 Hz is also exhibiting a slight lowering in level. However, the highest frequency used in this project is 8000 Hz (in the dead regions' test) and this deviation at the 10000 Hz 1/3 octave band should not have any effect whatsoever on this 8000 Hz stimulus or any other stimuli used.

When the response of the pink noise was adjusted, a measurement of the different stimuli was made. It was earlier decided that, since the spectrum levels of the babble

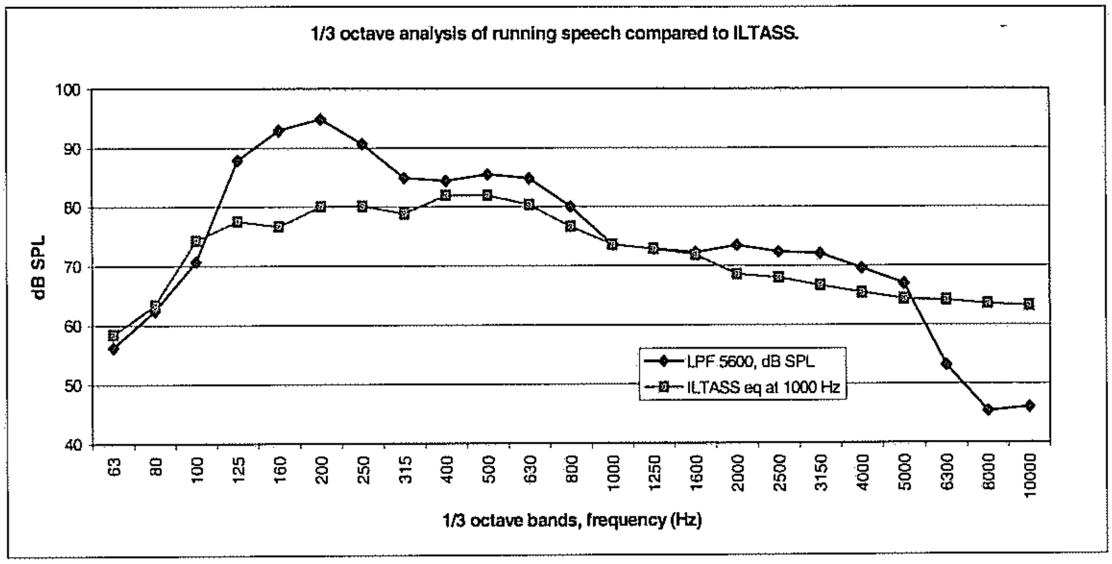

Figure 6. 1/3 octave analysis of babble noise in the broadband filter condition, LP 5600 Hz compared with ILTASS prior to adjustment. The right attenuator was set to 99.9, the left to 20.0, and the amplifier at level 3.The squares denote the ILTASS (Byrne et al., 1994). The diamonds denote the running speech in the broadband filtered condition, LP 5600 Hz, measured through all the equipment and the earphones using a 6-cc coupler and a weight of 455 grams. The levels are equalized at 1000 Hz.

noise were slightly higher in the high frequencies than the ILTASS, the speech should match the babble. Now however, reconsidering the size of the deviant levels was decided that a correction was required. In Figure 6 the results from the comparison between the broadband filter condition, LP 5600 Hz, and ILTASS are presented.

The difference between the running speech in the broadband condition and the ILTASS was calculated. For the lower frequency bands minor correction values

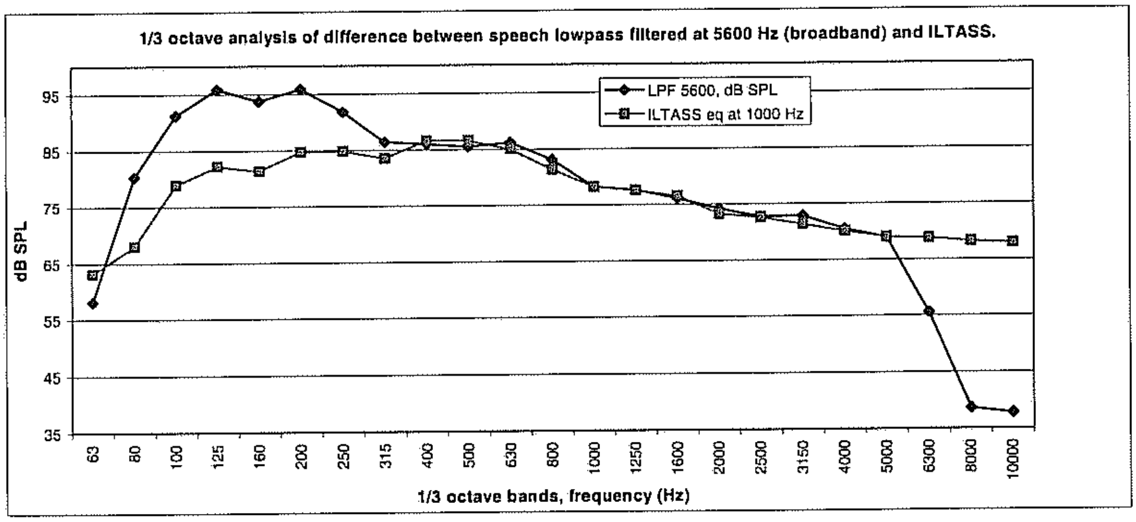

Figure 7. 1/3 octave analysis of babble noise in the broadband filter condition, LP 5600 Hz, compared with ILTASS after adjustment. The right attenuator was set to 99.9, the left to 20.0, and the amplifier at level 3.The squares denote the ILTASS (Byrne et al., 1994). The diamonds denote the running speech in the broadband filtered condition, LP 5600 Hz, measured through all the equipment and the earphones using a 6-cc coupler. The levels are equalized at 1000 Hz.

(ranging from about 1 dB to 5 dB) were calculated, which means that the correction values were substantially smaller than the actual difference. This was due to the discrepancy between the response using the 6-cc coupler or the KEMAR, as mentioned earlier.

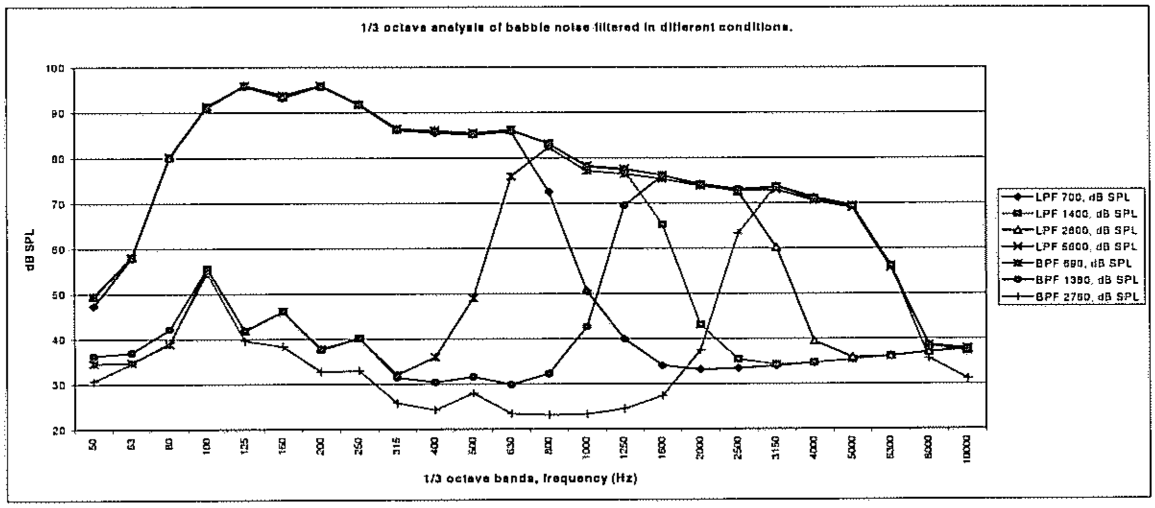

Figure 8. 1/3 octave analysis of running speech in the different filtered conditions used measured through all the equipment and the earphones using a 6-cc coupler and a weight of 455 grams. The right attenuator was set to 99.9, the left to 20.0, and the amplifier at level 3. The diamonds denote the running speech in the low-pass filtered condition 700 Hz. The squares denote the running speech in the low-pass filtered condition 1400 Hz. The triangles denote the running speech in the low-pass filtered condition 2800 Hz. The crosses denote the running speech in the broadband condition 5600 Hz. The stars denote the running speech in the high-pass filtered condition 690 Hz. The filled circles denote the running speech in the high-pass filtered condition 1380 Hz. The vertical lines denote the running speech in the high-pass filtered condition 2760 Hz. All filters have their cut-off frequency at 5600 Hz.

2.5.5 Final calibration measurement

An adjustment using the equalizer was made to match the speech and the babble to ILTASS. In Figure 6 are the results from the measurement before and after the adjustment to ILTASS shown. No difference is larger than 1.8 dB and on average 0.3 dB, which is reasonable good. The exceptions are in the low frequency bands 50-315 Hz and in the high frequencies where of course the filter is cutting. As mentioned above, the difference in the lower frequencies are due to it being measured in a 6 cc-coupler and their acquittal has already been taken into account for.

Finally, the responses of all the different stimuli were measured using the 6-cc coupler. Measurements had been made for all the stimuli during the previous calibration measurements. They are not accounted for here. The relevant cases have been presented in the sections above. A few examples of the results from these measurements are presented in Figures 8-10 to give a view of the stimuli used. The absolute values in dB SPL derived are not presented in tables due to their extensive size. The values measured are though the basis for the presentation of the results and are accounted for as required in chapter 3, Results.

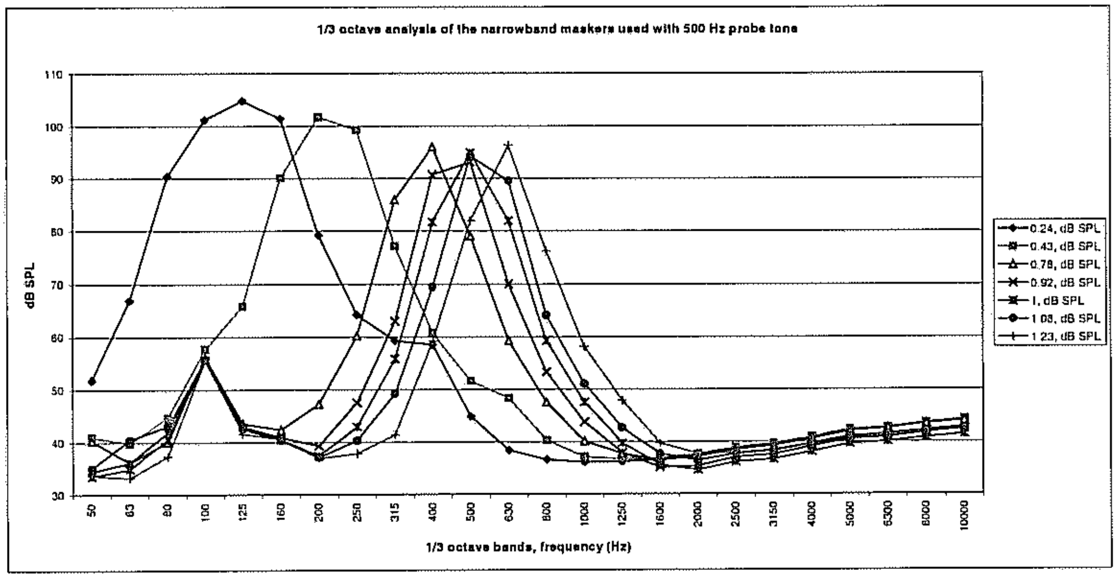

Figure 9. 1/3 octave analysis of the narrowband maskers used with 500 Hz probe tone measured through all the equipment and the earphones using a 6-cc coupler and a weight of 455 grams. The right attenuator was set to 20.0, the left to 99.9, and the amplifier at level 3. The diamonds denote the narrow band filter with a center frequency of 0.24 times the probe tone. The squares denote the narrow band filter with a center frequency of 0.43 times the probe tone. The triangles denote the narrow band filter with a center frequency of 0.78 times the probe tone. The crosses denote the narrow band filter with a center frequency of 0.92 times the probe tone. The stars denote the narrow band filter with a center frequency of 1 times the probe tone. The filled circles denote the narrow band filter with a center frequency of 1.08 times the probe tone. The vertical lines denote the narrow band filter with a center frequency of 1.23 times the probe tone.

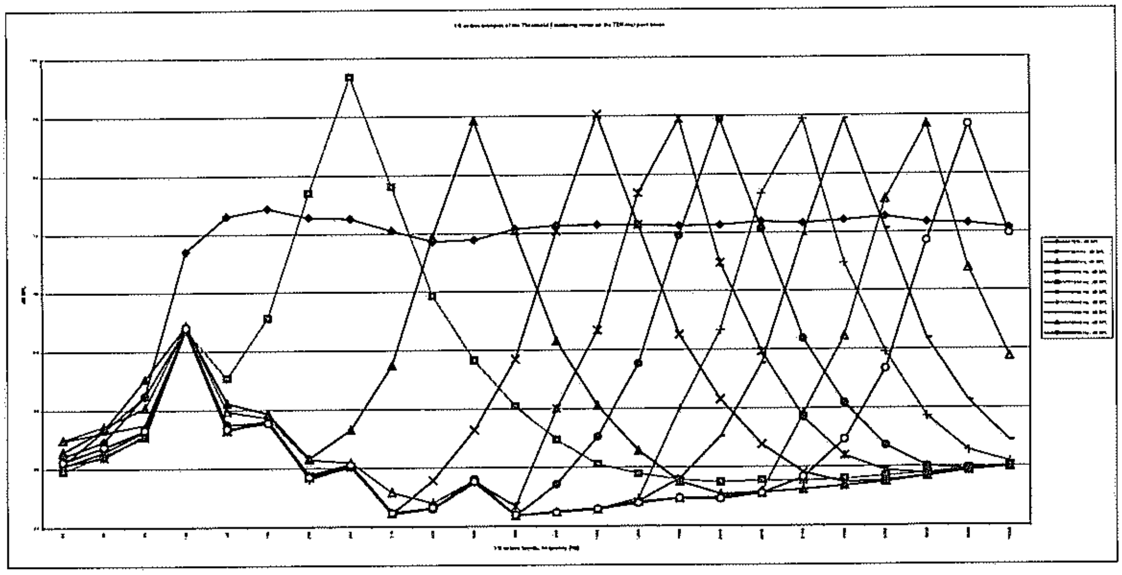

Figure 10. 1/3 octave analysis of the dead regions' test. All the tones and the TEN noise are measured through all the equipment and the earphones using a 6-cc coupler and a weight of 455 grams. For the TEN noise was the right attenuator was set to 40.0, the left to 99.9, and the amplifier at level 3. For the tones were the reverse settings on the attenuators used. The diamonds denote TEN noise. The squares denote the 250 Hz tone. The filled triangles denote the 500 Hz tone. The crosses denote the 1000 Hz tone. The stars denote the 1500 Hz tone. The filled circles denote the 2000 Hz tone. The vertical lines denote the 3000 Hz tone. The horizontal lines denote the 4000 Hz tone. The triangles denote the 6000 Hz tone. The circles denote the 8000 Hz tone.

2.5.6 Summary of calibration

The results of the first calibration indicated that the earphones had a non-flat response. Due to this another measurement of the earphones using KEMAR was made to examine if the external ear canal would make the frequency response flat. This was not the case. The equalizer was set to adjust the deviant levels using the difference between the CD direct response and the 6-cc coupler. A second equalizer was added to enable the shaping of the spectrum according to the POGO II. Another set of calibration measurements was made. As mentioned above, the speech stimuli were filtered to match the babble noise, which differed from the ILTASS to some extent. Since the difference from ILTASS was larger in the higher frequencies than in the lower, it was now decided that the equalizer should also adjust for this difference. The final calibration was made after this adjustment to determine the absolute levels of the stimuli.

3. RESULTS

3.1 Transiently Evoked Otoacoustic Emissions (TEOAE)

Below are the results from the measurement of TEOAE presented in Table 5. The TEOAE was used only to register the function of the subjects' outer hair cells. The CES value in dB SPL was used for this evaluation. This is a simple method of determine both the presence and likelihood of hearing loss for the subjects at this stage of the project rather than looking at individual bands, since only normal hearing subjects were tested. However, the emission strength, waveform reproducibility, and signal-to-noise ratio for five frequency bands were also recorded along with an overall value for each of these parameters. These values are not regarded here but will probably be included when the comparison between normal hearing subjects and hearing impaired can be done.

Table 5. Coherent Emission Strength (CES) values in dB SPL for each of the subjects.

Subject	dB SPL
1	9
2	5
3	15,4
4	6,1
5	11,9

The finding is that all of the subjects have a CES value that is well above 0 dB, on average about 9.5 dB SPL and ranging between 5 dB SPL and 15.4 dB SPL.

3.2 Dead regions' test

The results from the dead regions' test are presented for each subject in Figures 12 to 16. In Figure 11 the mean results and standard deviations are presented.

Generally speaking, all of the subjects display thresholds in quiet below 33 dB SPL. A threshold at 33 dB SPL roughly equals a threshold at 10 dB HL, if the ISO 389-1 (ISO, 1998) dB RETSPL is used. However, the RETSPL value differ between different frequencies, e.g. it is 25.5 at 250 Hz, 7 at 1000 Hz, and 13 at 8000 Hz. If this is taken into account there is still no threshold in dB HL that is over 20 dB HL. Although, it is necessary to remember that this standard is designed for mainly TDH 39 earphones and not the ones used in this test. But it gives an idea of if the thresholds are reasonable for normal hearing subjects.

The average threshold in TEN lies at 70.6 dB SPL, which almost was the level of the noise used. However, while the standard deviations in the frequencies 250 Hz to 3000 Hz are small, about 3.6 dB on average, they are quite large in the higher frequencies 4000 Hz to 8000 Hz, about 11.6 dB on average. The correlation coefficient between the thresholds in quiet and in noise is 0.71.

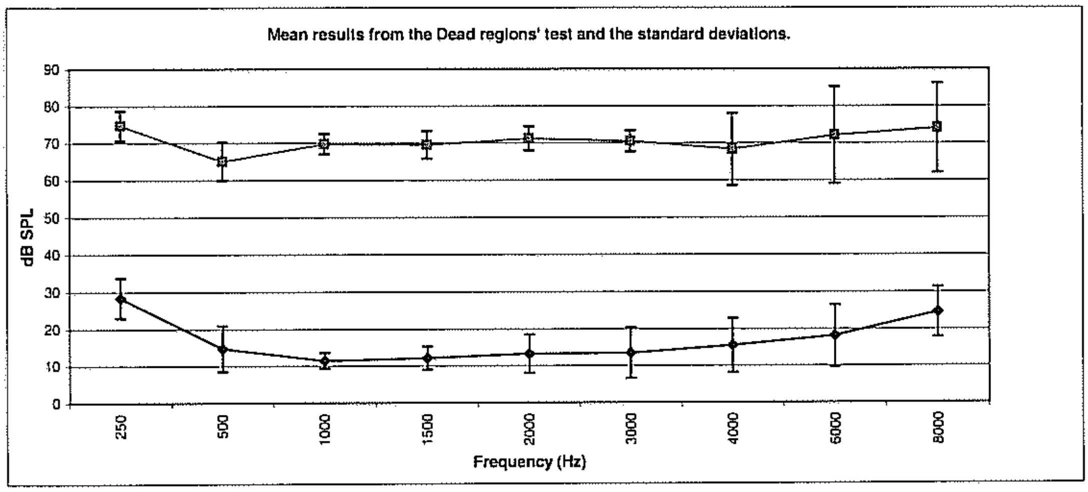

Figure 11. The average thresholds in the Dead regions' test and the standard deviations. The squares denote the average thresholds in quiet. The diamonds denote the average thresholds in TEN. The vertical bars represents the standard deviations derived separately for the thresholds in quiet and in TEN. N=5.

Subject 1 displays thresholds in quiet that on average are around 10 dB SPL. Two frequencies, 250 Hz and 6000 Hz, have higher thresholds than the average for this subject, with 9.4 dB and 7.2 dB respectively. Three thresholds lie outside the 95% confidence interval, 500 Hz, 4000 Hz, and 8000 Hz. The average threshold in TEN for subject 1 is 65.4 dB SPL, which is almost 5 dB lower than the level of the TEN. The most notable feature in the thresholds in noise is the thresholds at 500 Hz, which is 13.8 dB lower than the TEN and 9.2 dB lower than the subjects' average. And all of the thresholds in noise lie within the 95% confidence interval except the 500 Hz tone.

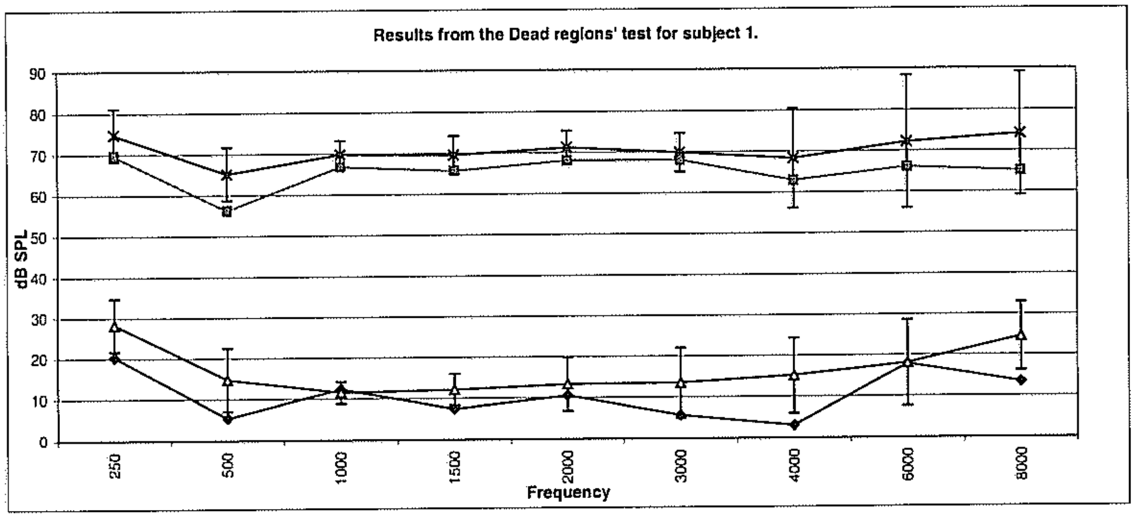

Figure 12. Thresholds in quiet and in TEN using the Dead regions' test for subject 1. The squares denote the thresholds in quiet and the diamonds denote the thresholds in TEN. The values are the mean of 3-5 repetitions. The triangles represent the mean thresholds in quiet and the crosses the mean thresholds in noise, n=5. The vertical bars represents the 95% confidence interval derived separately for the thresholds in quiet and in TEN

The correlation coefficient between the thresholds in quiet and in noise for this subject is 0.53.

Subject 2 has an average threshold in quiet at 19.5 dB SPL. The thresholds differ quite a lot between frequencies. Three frequencies 250 Hz, 6000 Hz, and 8000 Hz have higher thresholds than the average at, in the same order, 32.8 dB SPL, 26.3 dB SPL, and 31.2 dB SPL. That is a difference from the subjects' average threshold by 13.3 dB, 6.8 dB, and 11.7 dB respectively. Three frequencies, 500 Hz, 1000 Hz, and 1500 Hz have about 6.5 dB lower thresholds than the average. The threshold at 2000 Hz is even lower with a difference from the average by 13.9 dB. All of the thresholds lie within the 95% confidence interval except the threshold at 2000 Hz. The average threshold in TEN for subject 2 is 78.6 dB SPL, which is 8.6 dB higher than the noise. What catches the eye, are the thresholds at the frequencies 250 Hz, 4000 Hz, 6000 Hz, and 8000 Hz, which are at a level of 9.4 dB, 10.8 dB, 22.7 dB, and 22.8 dB higher than the noise. The two highest frequencies are really remarkable because regarding the average threshold of the subject, 78.6 dB SPL, they are still more than 10 dB higher. Furthermore, they lie outside two standard deviations and outside the 95% confidence interval along with the 4000 Hz tone. The correlation coefficient between the thresholds in quiet and in noise for the subject 2 is 0.70.

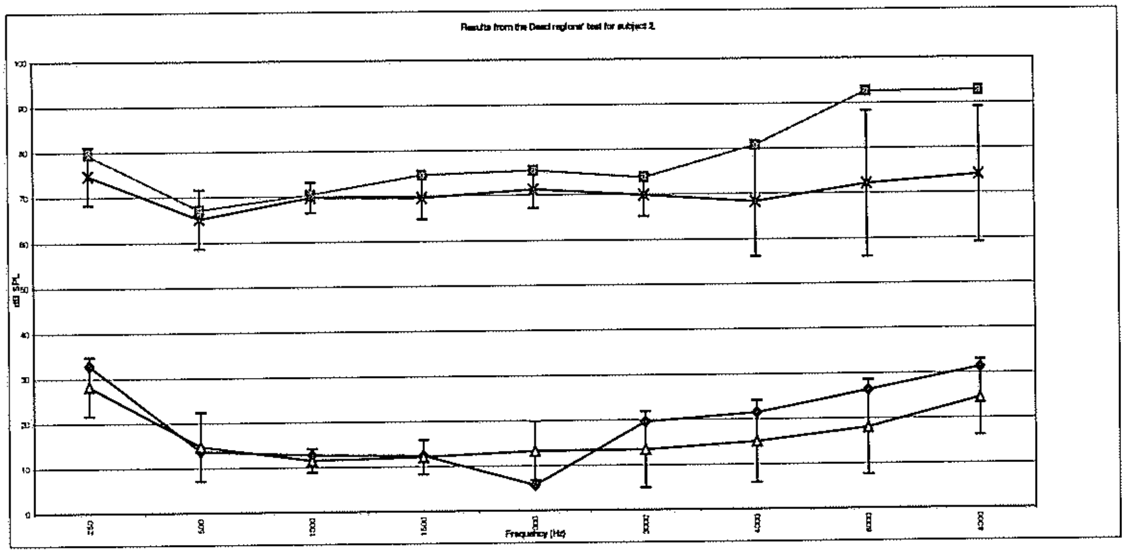

Figure 13. Thresholds in quiet and in TEN using the Dead regions' test for subject 2. The squares denote the thresholds in quiet and the diamonds denote the thresholds in TEN. The values are the mean of 3-5 repetitions. The triangles represent the mean thresholds in quiet and the crosses the mean thresholds in noise, n=5. The vertical bars represents the 95% confidence interval derived separately for the thresholds in quiet and in TEN

Subject 3 displays a threshold in quiet that on average is 16.2 dB SPL. Two frequencies, 250 Hz and 8000 Hz, have almost about 10 dB higher thresholds than the average for this subject. The average threshold in TEN is 73.4 dB SPL, which is about 3 dB higher than the noise. The only notable feature of the thresholds in TEN is the threshold at 8000 Hz, which has a level that is 9.8 dB higher than the noise, but is

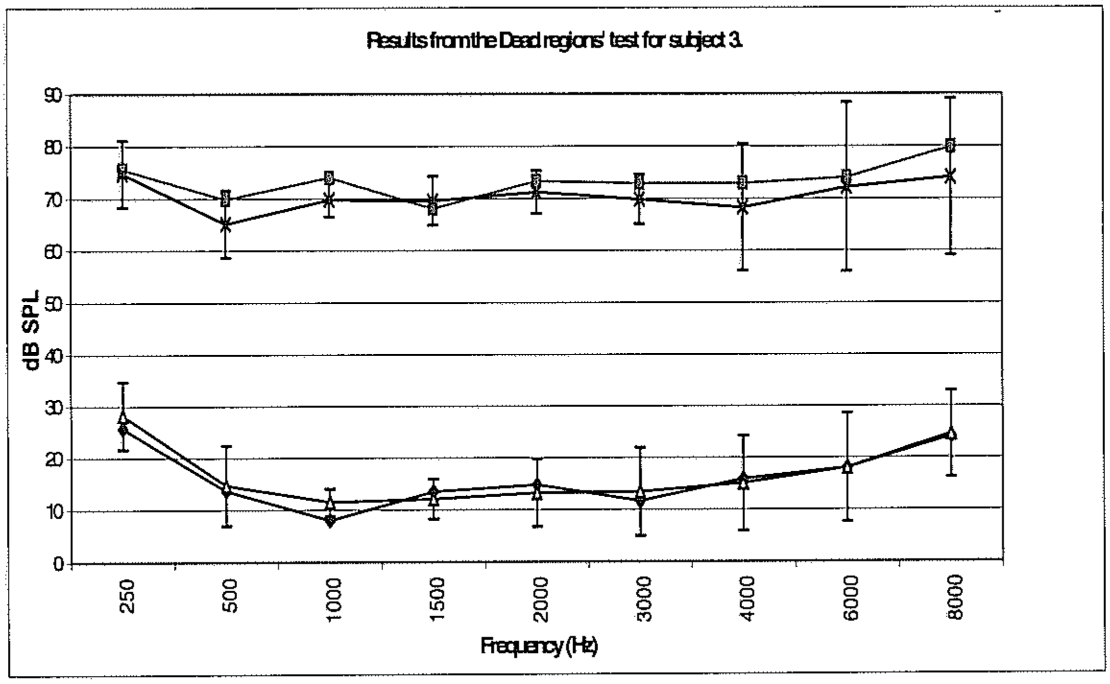

Figure 14. Thresholds in quiet and in TEN using the Dead regions' test for subject 3. The squares denote the thresholds in quiet and the diamonds denote the thresholds in TEN. The values are the mean of 3-5 repetitions. The triangles represent the mean thresholds in quiet and the crosses the mean thresholds in noise, n=5. The vertical bars represents the 95% confidence interval derived separately for the thresholds in quiet and in TEN

only 6.4 dB higher than the subjects' average threshold in noise. All of the thresholds lie within the 95% confidence interval except the one at 1000 Hz in noise. The correlation coefficient between the thresholds in quiet and in noise for the subject is 0.61.

The average threshold in quiet for subject 4 is at 20.1 dB SPL. It is important to observe that this average is calculated over the frequencies 500 Hz to 8000 Hz, thus omitting the 250 Hz tone. This was done since the subject was unable to establish a threshold in the TEN at that frequency and it seems more appropriate to account for the frequencies obtained in both conditions. Two frequencies, 1000 Hz and 8000 Hz, deviate markedly from the average threshold of the subject by -7.2 dB and 7.6 dB respectively. The average threshold in TEN is accordingly calculated without taking the 250 Hz tone into account. The average threshold in TEN for subject 4 is 69.3 dB SPL, which is almost spot on the level of the noise. The fact that the subject could not

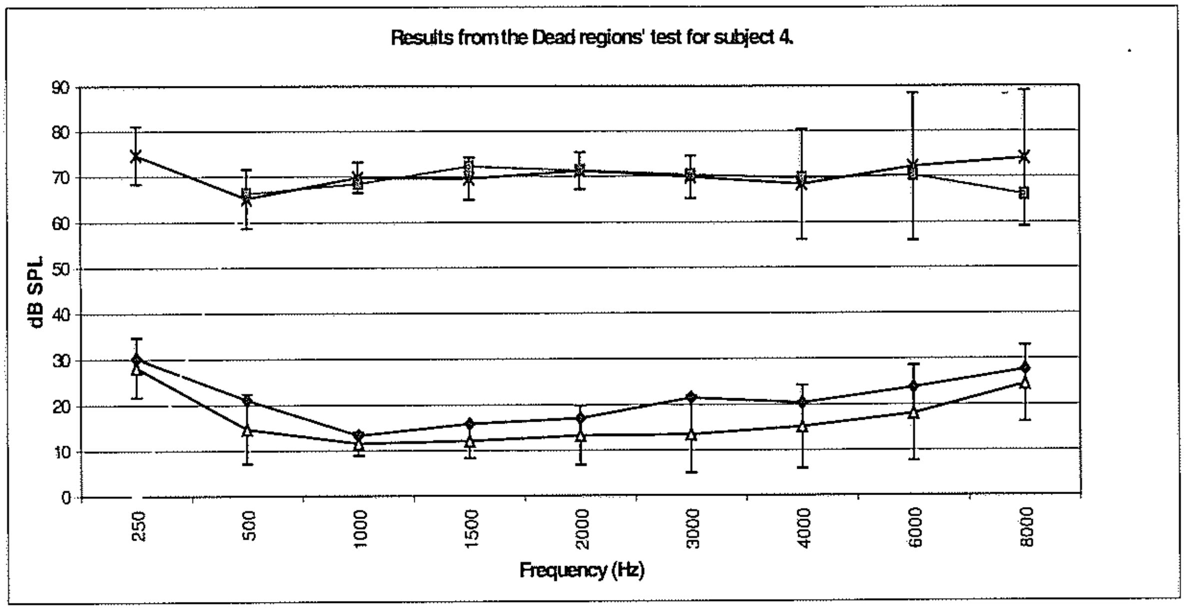

Figure 15. Thresholds in quiet and in TEN using the Dead regions' test for subject 4. The squares denote the thresholds in quiet and the diamonds denote the thresholds in TEN. The values are the mean of 3-5 repetitions. The triangles represent the mean thresholds in quiet and the crosses the mean thresholds in noise, n=5. The vertical bars represents the 95% confidence interval derived separately for the thresholds in quiet and in TEN

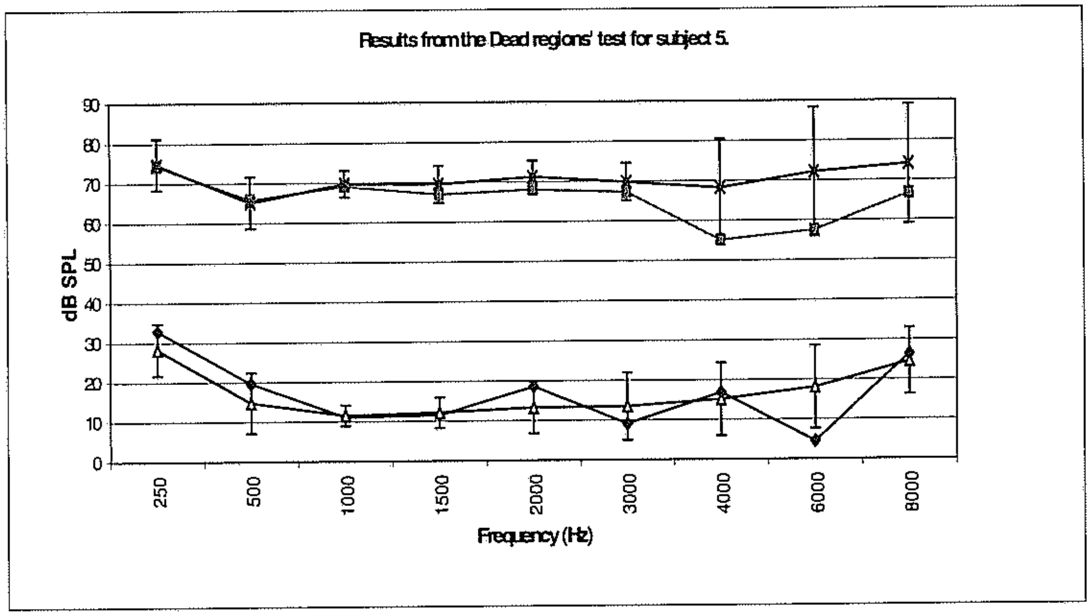

Figure 16. Thresholds in quiet and in TEN using the Dead regions' test for subject 5. The squares denote the thresholds in quiet and the diamonds denote the thresholds in TEN. The values are the mean of 3-5 repetitions. The triangles represent the mean thresholds in quiet and the crosses the mean thresholds in noise, n=5. The vertical bars represents the 95% confidence interval derived separately for the thresholds in quiet and in TEN

establish a threshold at 250 Hz is interesting. The subject commented on this that the level of the tone was decreasing until a certain level and them it lay steadily at that level with further decrease on the attenuator. This frequency was later re-tested at second session with the same outcome. All of the thresholds lie within the 95%

confidence interval. The correlation coefficient between the threshold in quiet and in noise for subject 4 is -0.49.

Subject 5 displays a threshold in quiet that on average is 16.7 dB SPL. Two frequencies, 250 Hz and 8000 Hz, have 16.2 dB and 9.4 dB higher thresholds than the subjects' average. The average threshold in TEN is 65.7 dB SPL, which is 4.3 dB lower than the noise level. Three frequencies, 250 Hz, 4000 Hz, and 8000 Hz, deviate notably from the average threshold with 8.6 dB, -10.6 dB, and 8.2 dB respectively and in the same order 4.3 dB, -14.9 dB, and -12.5 from the level of the TEN. All thresholds lie within the 95% confident interval except at 6000 Hz in quiet. The correlation coefficient between the thresholds in quiet and in noise for subject 5 is 0.51.

3.3 Psychophysical Tuning Curves

The results from the PTC-test are unanimous. All of the subjects have the lowest signal-to-noise ratio at the center frequencies that match the frequencies of the probe tones. The level of the different maskers increases when the distance in frequency increases in either direction from the probe tones used. No deviations from this pattern can be detected.

There is no need to examine the results for each PTC for each subject in detail, but the results for one subject, subject 2, will serve as an example. The results are presented in figure 17

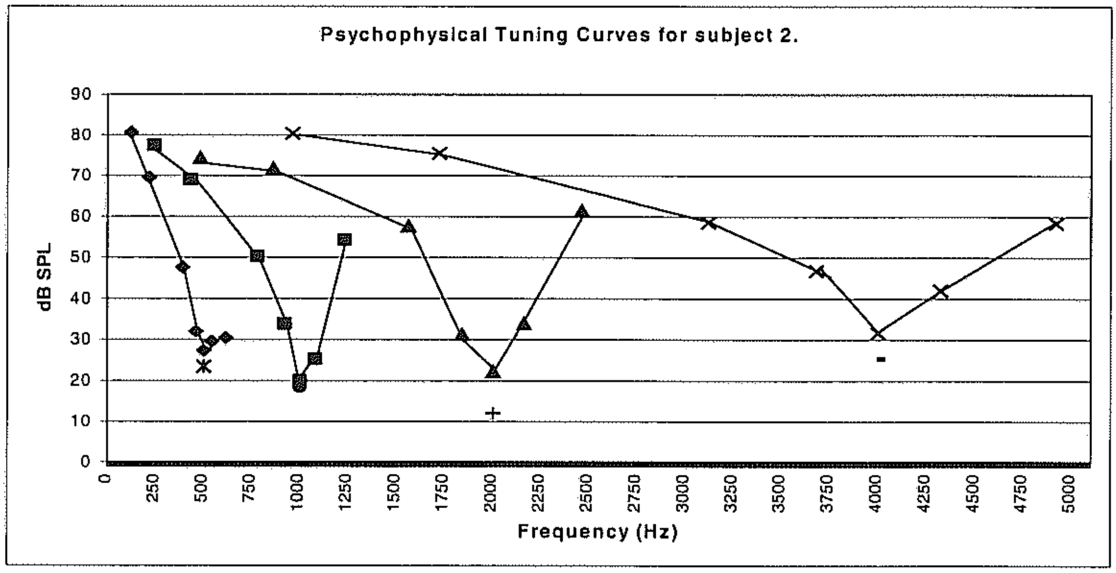

Figure 17. Psychophysical tuning curves (PTC) for subject 2. The diamonds denote maskers used with the 500 Hz probe tone. The squares denote the maskers used with the 1000 Hz probe tone. The filled triangles denote the maskers used with the 2000 Hz probe tone. The crosses denote maskers used with the 4000 Hz probe tone. The star denotes the 500 Hz probe tone. The filled circle denotes the 1000 Hz probe tone. The plus denotes the 2000 Hz probe tone. The horizontal line denotes the 4000 Hz probe tone.

When the 4000 Hz probe tone is used at a level of 10 dB SL, which equals 25.3 dB SPL, the masker with the lowest center frequency, 920 Hz, has to be 80.4 dB SPL to just mask the probe tone. The masker with the second, third, and fourth lowest center frequencies, 1720 Hz, 3120 Hz, and 3680 Hz respectively, have to be, in the same order, 75,5 dB SPL, 58,6 dB SPL, and 46,7 dB SPL to just mask the probe tone. The masker with the same center frequency as the probe tone only has to reach a level of 31.6 dB SPL to sufficiently mask the probe tone. The maskers with the two highest center frequencies, 4320 Hz and 4920 Hz, have to be 42,0 dB SPL and 58,4 dB SPL respectively to just mask the probe tone. Thus, the lowest signal-to-noise ratio is found when the masker with the same center frequency as the probe tone is used.

3.4 CUNY-sentences and VCV-utterances

In Figures 16-19 the results from the speech are presented. Observe that it is only the results obtained from one tested individual, subject 4, at a single session. This means that the results are not significant, but they serve as an indication of how the tests are working in comparison with the intentions of the test construction.

3.4.1 CUNY-sentence test

The CUNY-lists were presented at a level of 82.8 dB SPL in the QH condition, 40.7 dB SPL in the QL, and 82.8 dB SPL for the speech stimuli and 83.5 dB SPL for the babble noise in the N condition, a signal-to noise ratio of -0.7 dB.

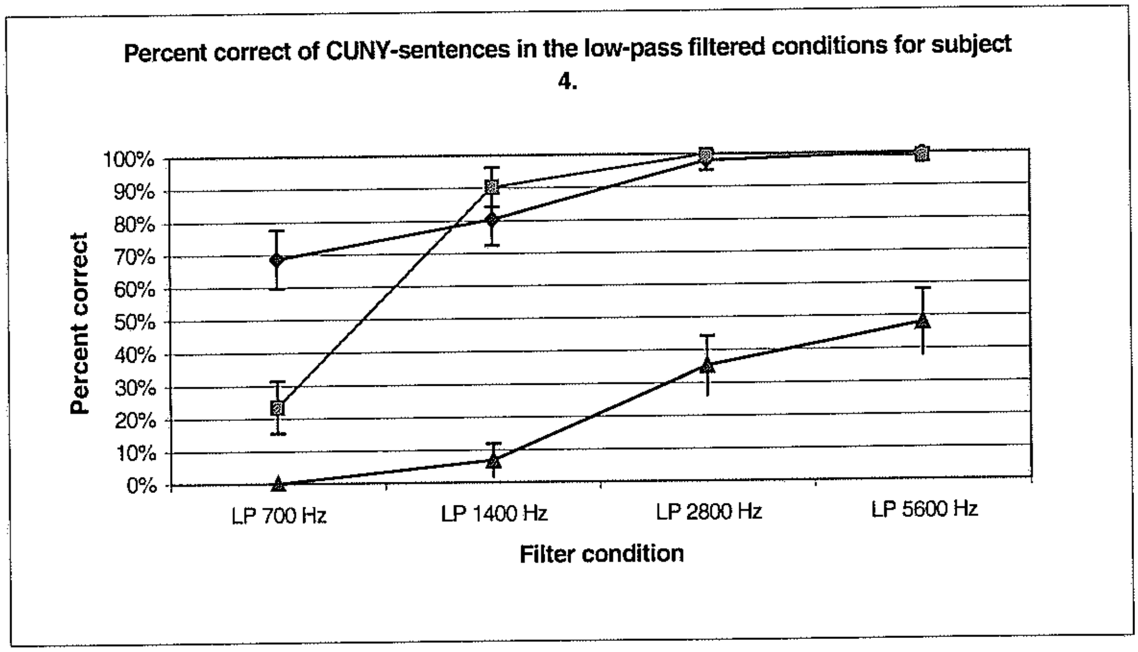

Figure 18. Scores correct for subject 4 in the low-pass filtered conditions using the CUNY-lists. The diamonds denote the quiet high (QH) condition, which means an overall level of 82.8 dB SPL in the broadband LP 5600 filer condition. The squares denote the quiet low (QL) condition, which means an overall level of 40.7 dB SPL in the broadband LP 5600 filer condition. The triangles denote the noise (N) condition, which means an overall level of 82.8 dB SPL of the speech and 83.5 dB SPL of the babble noise in the broadband LP 5600 filer condition, thus a signal-to-noise ratio of -0.7 dB. The vertical bars represents the 95% confidence interval.

As seen in Figure 16 over the percent correct of the CUNY's in the low-pass filtered conditions, the scores correct for the condition QH are 69% in the LP 700, 80% in the LP 1400, 98% in the LP 2800, and 100% in the broadband LP 5600. In the condition QL the scores correct are 24% in the LP 700, 90% in the LP 1400, 100% in the LP 2800, and 99% in the LP 5600. The N condition results in scores correct that are 0% in the LP 700, 7% in the LP 1400, 35% in the LP 2800, and 48% in the LP 5600. Thus, the scores correct are increasing with increasing cut-off frequency in the filter. The only exception is the LP 5600 in the condition QL that is 1% lower than the results obtained in the LP 2800 in the same condition. 1% equals about 1 word of a total of 102 in a CUNY-list. It is notable that the subject was not able to understand one single word in the N condition using the LP 700.

In Figure 17 the scores correct are presented for the CUNY's in the band-pass filtered conditions. This graph also includes the broadband filtered condition, LP 5600. The scores correct for the condition QH are as follows. The results are 83% scores correct in the BP 2760 and 100% for both the BP 1360 and BP 690. In the QL condition the scores correct are 10% in the BP 2760 and 100% in both the BP 1380 and BP 690. In the N condition the results are 0% correct in both BP 2760 and BP 1380 and only 9% in the BP 690. In the broader conditions, BP 1380 and BP 690 the subject is able to understand all of the stimuli in both the QH and QL condition. There is a substantial difference between the percent correct in the most narrowly filtered condition, BP2760, where the subject can recognize 83% correct in the QH but only 10% in the

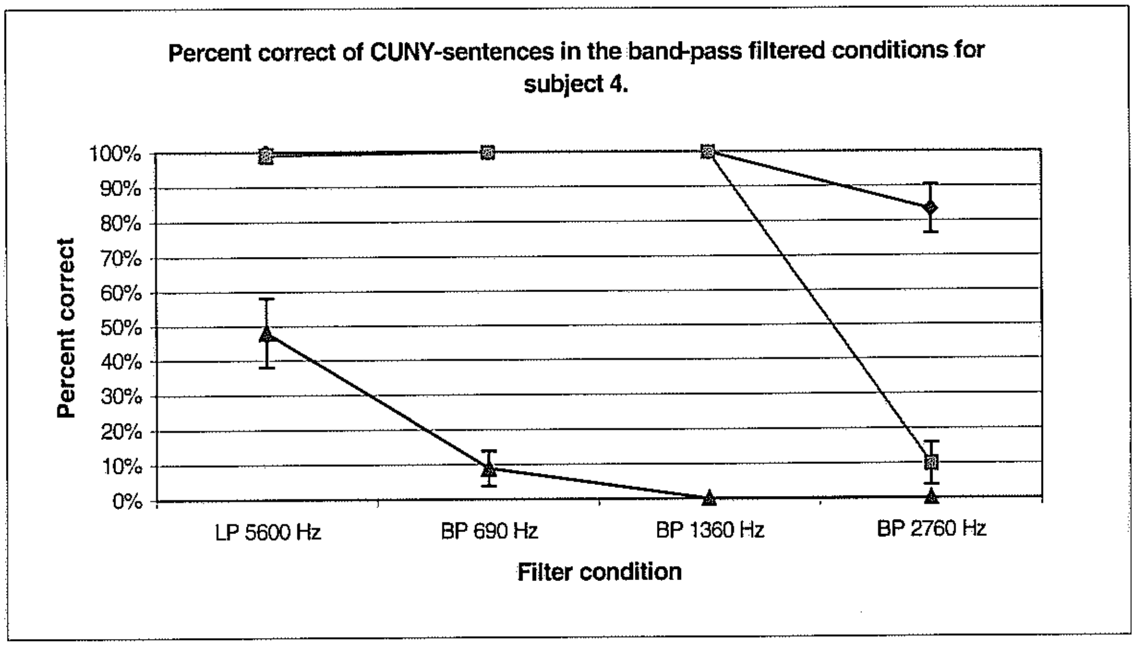

Figure 19. Scores correct for subject 4 in the band-pass filtered conditions using the CUNY-lists. The diamonds denote the quiet high (QH) condition, which means an overall level of 82.8 dB SPL in the broadband LP 5600 filer condition. The squares denote the quiet low (QL) condition, which means an overall level of 40.7 dB SPL in the broadband LP 5600 filer condition. The triangles denote the noise (N) condition, which means an overall level of 82.8 dB SPL of the speech and 83.5 dB SPL of the babble noise in the broadband LP 5600 filer condition, thus a signal-to-noise ratio of -0.7 dB. The vertical bars represents the 95% confidence interval.

QL. In the N condition the subject managed only to recognize 9% correct in the broadest filter condition, BP 690. There were 0% correct in the two narrower conditions, BP 1380 and BP 2760.

Confidence intervals were calculated using the binomial theorem. It should be noted that this calculation provides a good approximation for values between 15% to 85%. Values outside these should be calculated using arc-sine transform. It can be said that the small spread of the 95% confidence intervals indicates a consistency in the results obtained. In the results the confidence interval is ranging from 8% till 10 % in the scores correct that lie within the 15% to 85% interval where the approximation is valid, i.e. for the LP 700 in the conditions QH and QL, the LP 1400 in QH, the LP 2800 and the LP 5600 in N, and the BP 2760 in QH. If the correlation coefficients are calculated between the results from all the filter conditions in QH and QL, QH and N, and QL and N, they are 0.77, 0.53, and 0.51 respectively.

3.4.2 VCV's

The VCV-lists were presented at a level of 87.3 dB SPL in the QH condition, 45.9 dB SPL in the QL, and 87.3 dB SPL for the speech stimuli and also for the babble noise in the N conditions rendering a signal-to-noise ratio of ±0 dB.

In figure 18 are the scores correct of the VCV's presented as a function of the low-pass filters used. In the QH condition the scores correct are 33% in the LP 700, 57%

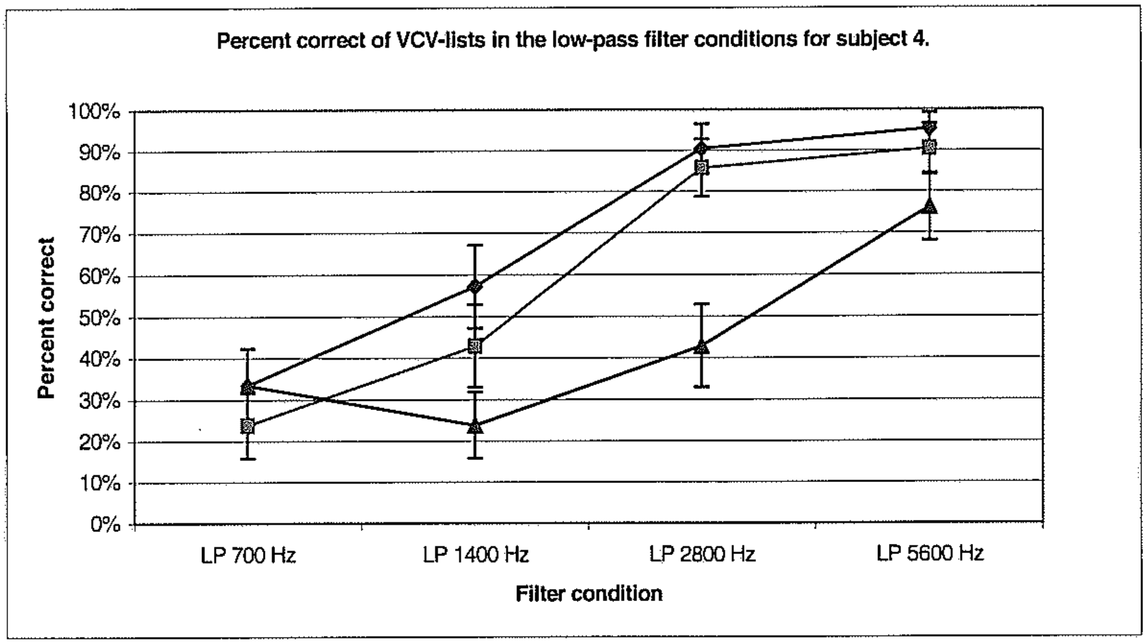

Figure 20. Scores correct for subject 4 in the low-pass filtered conditions using the VCV-lists. The diamonds denote the quiet high (QH) condition, which means an overall level of 87.3 dB SPL in the broadband LP 5600 filer condition. The squares denote the quiet low (QL) condition, which means an overall level of 45.9 dB SPL in the broadband LP 5600 filer condition. The triangles denote the noise (N) condition, which means an overall level of 87.3 dB SPL of the speech and 87.3 dB SPL of the babble noise in the broadband LP 5600 filer condition, thus a signal-to-noise ratio of ±0 dB. The vertical bars represents the 95% confidence interval.

in the LP 1400, 90% in the LP 2800, and 95% in the broadband LP 5600. In the QL the scores correct are 24% in the LP 700, 43% in the LP1400, 86% in the LP 2800, and 90% in the LP 5600. In the N condition the subject has the following scores correct. In the LP 700 the score correct is 33%, but in the LP 1400 the score is only 24%. The scores correct are 43% and 76% respectively in the LP 2800 and LP 5600. The scores correct are increasing with increasing cut-off frequency for the low-pass filtered VCV's as in the CUNY's. There is however an exception. The score obtained in the N condition using the LP 1400 is lower, 24%, than for the filter condition LP 700, which is 33%.

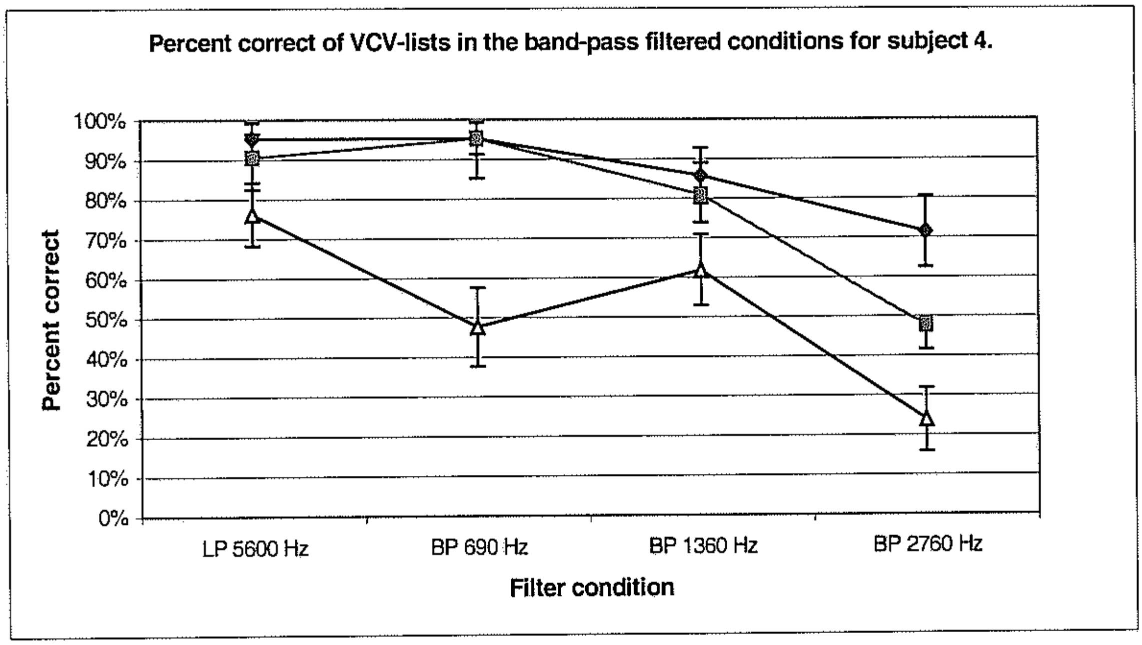

Figure 21. Scores correct for subject 4 in the band-pass filtered conditions using the VCV-lists. The diamonds denote the quiet high (QH) condition, which means an overall level of 87.3 dB SPL in the broadband LP 5600 filer condition. The squares denote the quiet low (QL) condition, which means an overall level of 45.9 dB SPL in the broadband LP 5600 filer condition. The triangles denote the noise (N) condition, which means an overall level of 87.3 dB SPL of the speech and 87.3 dB SPL of the babble noise in the broadband LP 5600 filer condition, thus a signal-to-noise ratio of ±0 dB. The vertical bars represents the 95% confidence interval.

The results from the band-pass filtered conditions using the VCV's are presented in Figure 19. In the QH condition the scores correct are 71% in the BP 2760, 86% in the BP 1380, and 95% in the BP 690. In the QL condition the scores correct are 48% in the BP 2760, 81% in the BP 1380, and 95% in the BP 690. The results of the scores correct are in the N condition 24% in the BP 2760, 62% in the BP 1380, and 48% in the BP 690. It should be noted that the BP 1360 has a higher score than the wider BP 690. The pattern is once again repeating itself. The speech scores are increasing with greater bandwidth of the filters, just as it did in the scores from the low-pass and band-passed filtered CUNY's and the low-pass filtered VCV's with a few exceptions.

Confidence intervals were calculated using the binomial theorem with its valid approximation for values between 15% to 85%. In the results the confidence interval is ranging from 8% till 10 % in the scores correct that lie within the 15% to 85% interval, that is for the LP 700 and the LP 1400 in the conditions QH and QL, the BP

2760 in QH, the BP 1360 and the BP 2760 in QL, and in all conditions using the band-pass filtered VCV's in N. Once again it can be said that the small spread of the 95% confidence intervals indicates a consistency in the results.

The correlation coefficients between the results from all the filter conditions in QH and QL, QH and N, and QL and N, are 0.97, 0.65, and 0.74 respectively.

3.4.3 CUNY's and VCV's

Here are a brief encounter of the correlation between the results obtained using the CUNY-lists and the VCV-lists. The reader is exhorted that these figures are not significant but rather illustrative.

The correlation between the results from the QH condition in the different filtered conditions using CUNY's and the VCV's is 0.98. The correlation between the results from the QL condition in the different filtered conditions using CUNY's and the VCV's is 0.77. The correlation between the results from the N condition in the different filtered conditions using CUNY's and the VCV's is 0.60.

4. DISCUSSION AND CONCLUSIONS

4.1 TEOAE

The results from the measurement of the transiently evoked otoacoustic emissions indicate normal outer hair cell functions for all of the subjects. 0 dB is the estimated level at which mild hearing loss might be present or likely to occur (LePage et al., 2001). No CES values are below 0 dB and it can thus be concluded that the subjects are displaying normal outer hair cell function.

4.2 Dead regions' test

The average results from the measurements in the TEN are close to the level of the TEN. This means that on average the noise does what it is supposed to do, i.e. to equalize the masked thresholds at roughly the same level as the TEN, in this case was 70 dB SPL, regardless of frequency tested. However, the standard deviation calculated for each individual frequency becomes increasingly larger in the higher frequencies, 4000 Hz to 8000 Hz, from being rather small in the lower frequencies, 250 Hz-3000 Hz. This indicates that there are generally larger differences from the masker level at 70 dB SPL at these frequencies.

The results of the thresholds obtained using the TEN for subject 1 display an average below the TEN at about 65 dB SPL. There is a large deviation, almost 14 dB lower, from the TEN at 500 Hz. The deviation also lies outside the 95% confidence interval, which indicates that it is not likely to believe that it is a random error. It seems as if the subject has the capability to handle a really bad signal-to-noise ratio at that frequency. However, the deviation in the noise resembles the one that is present in the quiet thresholds at the same frequency. The resemblance is to some extent verified by the value of the correlation coefficient. The correlation coefficient between the thresholds in quiet and in noise is 0.53, which is reasonably good.

Subject 3 has an average threshold in noise about 73 dB SPL, a little bit higher than the TEN. The thresholds are quite flat in the noise, but once again there seems to be a resemblance with the thresholds in quiet. There is a large deviation from the TEN at 8000 Hz. However, it does not exceed 10 dB and it lies within the 95% confidence interval for that frequency. Moore (e.g. 2001) states that a threshold that are at least 10 dB higher than the TEN indicates a dead region in that particular frequency region. The threshold that lies higher than the limit of confidence interval of 95% can not be explained as a random error and it is lower than the stipulated level of 10 dB higher than the TEN. The correlation coefficient between the thresholds in quiet and in noise is 0.61, which seems to verify the impression that the threshold in noise resembles the quiet one.

Subject 4 has almost flat thresholds in noise, average 69 dB, and they do not display any resemblance with the quiet ones. The calculation of the correlation ended up with a negative coefficient of -0.49, which is the lowest of all subjects. And accordingly, all of the thresholds lies within the 95% confidence interval. As mentioned in chapter 3, Results, the thresholds in quiet and in noise for 250 Hz have not been taken into account due to the fact that the subject was unable to establish the threshold in noise. The subject commented during testing that the tone decreased to a certain level and then lay steadily at the same level with further decrease on the attenuator. The Dead regions' test was repeated during a separate session and the same phenomenon occurred. It is likely to believe that the characteristics of the test can provide with an explanation for this. The noise was, as mentioned earlier, a broadband noise along with a set of pure tones. The pure tones were not pulsing. Since the frequency contents of the pure tone was also contained in the broadband noise it might be possible that a top-down triggering towards that specific frequency band in the noise was made. It is not likely that there was an actual tone present in due to equipment failure, since the linearity of all of the equipment had been thoroughly measured. More importantly, no one of the other subjects displayed such a deviation. However, a larger group of subjects need to be tested if any conclusions are to be drawn from this finding.

Subject 5 has thresholds in noise that are close to the TEN until 3000 Hz. At 4000 Hz there is a lowering in the threshold by about 15 dB and at 6000 Hz by about 13 dB. The latter one also lies outside the 95% confidence interval and can thus not be regarded as a random error but rather, as in the case with subject 1, an expression for high tolerance for unfavorable signal-to-noise ratio at this frequency. Once again there is a graphical resemblance between the thresholds in quiet and in noise that might account for these deviations. The correlation coefficient speaks for this to some extent. It is 0.51, which is acceptable.

The most interesting results are from subject 2. Subject 2 exhibits a deviation from the TEN of almost 23 dB in two frequencies, 6000 Hz and 8000 Hz, and almost 11 dB in one, 4000 Hz. According to Moore (2001) this indicates a dead region that starts from about 4000 Hz and pertains to the higher frequencies. To recollect, Moore (2001) states that a dead region is present if the TEN produces a rise in threshold that is at least 10 dB above the TEN when the TEN is presented at least at 10 dB SL. Since the highest probe tone frequency used in the PTC test was 4000 Hz, nothing can be said about the deviations in the two highest frequencies, 6000 Hz and 8000 Hz except that they are probably not a random error since they lie will outside the 95% confidence interval. However, regarding the results from subject 2's PTC test using the 4000 Hz probe tone we would thus expect that we would not find a characteristic "tip" in the PTC at the probe tone frequency, but we do indeed find such a "tip". But it still lies outside the confidence interval of 95%, which is interesting, but unexplainable at this stage of the project. We found that the correlation between the threshold in quiet and in noise for subject 2 was 0.70, which is rather a high correlation. Thus, the shape of the threshold in noise resembles the shape obtained in quiet.

The expected result using the TEN from the Dead regions' test would have been flat thresholds in noise for all subjects (Moore et al., 2000). However, for all subjects (except for subject 4) there are deviations from the expected thresholds. The differences in thresholds in noise might reflect the thresholds obtained in quiet. This was the case when correlation coefficients were calculated for each subject. On average, it might be possible to verify Moore's findings (Moore et al., 2000). The average thresholds are close to the TEN at 70 dB SPL, but the standard deviation is quite large in the frequencies 4000 Hz-8000 Hz. Looking at the average thresholds, there is still a resemblance between the thresholds in quiet and in noise, which is also the finding of the correlation coefficient between the two, 0.71, which is quite good. We would probably expect to see lower correlation coefficients for all the subjects individually and collectively, if the TEN equalized the thresholds at its presentation level.

It is hard to say that the Dead regions' test is working as it was meant to while observing the individual results, if we do not apply an interpretation of Moore et al. (Moore et al., 2000; Moore, 2001) that the by flat and equalized thresholds he means thresholds situated at the level of the TEN with the original shaping of the thresholds in quiet. However, this interpretation is not what Moore suggests (Moore et al., 2000; Moore, 2001). It is safe to say that the TEN is working sufficient on average but not on an individual basis.

4.3 Psychophysical Tuning Curves (PTC)

All subjects display PTC that have a "tip" at the same frequencies as the probe tones used, i.e. the lowest signal-to-noise ratio when the center frequency of the narrow-band masker is the same as the probe tone. This is the common finding from measurements of PTC (e.g. Moore et al., 2001, O'Loughlin, 1981, and Pickles, 1995) and thus the results for all subjects indicate a normal ability of frequency resolution.

It seems reasonably to conclude that the PTC-test is working sufficiently and as expected.

4.4 CUNY's and VCV's

Generally speaking, the overall finding is that speech scores increase with increasing bandwidth of the filters, i.e. the wider filter the better speech scores, which is the common finding for normal hearing subjects (e.g. Horwitz et al., 2002; Vickers et al., 2001). This seems to be the case for both the CUNY's and the VCV's in all conditions QH, QL, and N with a few exceptions. The application of the 95% confidence interval using the binomial theorem seems to indicate that the spread of the scores correct within the test is quite small. This speaks for the reliability of the speech tests. We are getting results that are reliable and seemingly without any larger random factors. This accounts for, as mentioned earlier, the scores correct that lie between 15% and 85%.

Subject 4 is the only subject tested using the speech tests and caution should be observed regarding the results, but the results display how the tests are working and if changes are required.

4.4.1 Noise (N) condition

In the CUNY-test there seems to be a problem in the noise condition when the speech filters are narrower, as in LP 700, BP 1380, and BP 2760. In all these conditions the results are 0% scores correct. This might be due to two reasons. Firstly, the signal-to-noise ratio was ±0 dB for the VCV's while it was -0.7 dB for the CUNY's. These signal-to-noise ratios were, if we recall, obtained using the broadband condition, LP 5600, and altering the level of the noise. This procedure was done for both the CUNY's and the VCV's. The aim was to obtain a signal-to-noise ratio where the speech scores were lowered by 1/3 compared to the QH. 2/3 of the scores correct was actually obtained using this method in the beginning of the session. It seems as if the signal-to-noise ratio for the CUNY's was working all right during this initial test, but was too hard for the subject when the "real" test in the broadband condition was conducted. One reason for this might be a question of concentration. Secondly, there is a difference between the speech tests used. The VCV-test is pretty straightforward. The task for the subject is to point out the consonant they are hearing on a chart containing all the consonants used. The subject was encouraged to guess when uncertain. The same exhortation was used testing the CUNY's. However, there are no guidelines in the CUNY-test like a chart, but rather the subject has to make up his own mind regarding what is said and repeat it aloud. This might have affected the reluctance to make a guess. Although, it appears more reasonable to assume that it is the slight difference in signal-to-noise ratio that has reduced the score. Furthermore it can be said that the overall level of the VCV-test was about 5 dB higher than for the CUNY-test. This has probably not affected the results since the signal-to-noise ratio would have been the same.

For the VCV's the noise condition is also providing two deviations. The first is found in the LP 700 condition where a higher score correct was obtained than in the broader LP 1400. It can not be established if this difference is statistically significant or not due to the fact that only a single subject has been tested. This might however be and a random error but the confidence interval speaks against that. The second one is that the BP 690 results in a higher score correct that the narrower BP 1380. Once again a random error is likely to have occurred, which, as in the former case, can not be established since these are results from only one subject. But still, the confidence interval indicates that it is not. A larger survey of subjects will hopefully provide an explanation for this discrepancy.

4.4.2 Quiet high (QH) condition

The QH condition for both the CUNY's and the VCV's in all the filtered conditions can presumably be concluded to be working as they were meant to be. For the CUNY's it was assumed that this condition, QH, would result in scores correct at or close to 100% for all filtered conditions with the exception for the most narrow, LP 700 and BP 2760. This seems to be the case. For the VCV's it was expected that larger differences between the filters would occur, since the subject has to rely on acoustical cues to a larger extent than in the CUNY-sentences where a semantic top-down process is likely to occur. This seems to be what has happened.

4.4.3 Quiet low (QL) condition

The real problem was the QL condition. A couple of tests were made using a level of presentation that was 1/3 of the level of the QH condition. At first this level was not allowed to drop below 6 dB SL and it was determined in the broadband condition, LP 5600, as mentioned earlier. At first this did not seem to cause a problem. However, when the LP 700, BP 1380, and BP 2760 were used the overall level was heavily reduced with the consequence that the stimuli dropped to or below the detection threshold for speech. Hence, this resulted in scores correct of 0%. To compensate for the overall loss of level for these filter conditions it was decided that 1/3 of the level of the QH should be used as long as it did not drop below 24 dB SL. This was still not sufficient for the filtered conditions BP 1380 and BP 2760. It was accordingly decided that a high frequency emphasis should be used even for the normal hearing subjects. The same model for spectral shaping, POGO II, was used for the normal hearing subjects as for the hearing impaired.

The results from subject 4 seem to imply that the overall level, 24 dB SL, probably is too high. The results obtained in the QL condition now resemble the ones in QH, which was not the aim of incorporating a QL level. The thought was rather that the scores correct for the QL condition would place themselves somewhere between the QH and the N condition. But still, there is a strong possibility that, if a lower overall level would be used, no results would have been obtained at all in the narrow filter conditions despite the applied high frequency emphasis. No conclusions can be deducted from this, but further testing is required to establish a more accurate level of presentation of the QL condition.

4.5 Summary

The underlying hypothesis of this thesis, that speech intelligibility in an individual can be more accurately predicted from measurements of frequency resolution and dead regions in the cochlea, can neither be confirmed nor rejected. The results and analysis of future tests using the equipment, stimuli, and procedures established in this thesis can hopefully answer that question.

The test of TEOAE is working as it was supposed to. The results provide a good estimation of the subjects' outer hair cell function. It also provides a complete record of waveform reproducibility, emission strength, and signal-to-noise ratio for separate frequency bands and overall as well. This may in a later stage of this project provide with insights into the effects of outer hair cell loss on speech intelligibility.

To draw conclusions from the results derived from the Dead regions' test the following can be said. The TEN does not seem to make the thresholds flat for normal hearing subjects. It rather just heightens the overall level and preserves the original threshold shaping obtained in the quiet condition. This is contrary to the intentions (Moore et al., 2000; Moore, 2001). Furthermore, the results from subject 2 indicate that the 10 dB limit, above which Moore (2001) states it is indicative of a dead region, may be to low.

The PTC-test seems to behave as it was intended to do. All subjects display "tips", i.e. the lowest signal-to-noise ratio, at the maskers with the same center frequencies as the probe tones, and the highest signal-to-noise ratio at the maskers with center frequencies that are most separated from the probe tone frequencies. This is the common finding for normal hearing subjects measuring PTC (e.g. Moore et al., 2001; O'Loughlin et al., 1981; Pickles, 1995).

The results from the speech tests for the only tested subject, subject 4, are indicating that the tests are generally behaving as they were supposed to. That is, for the three different conditions regarding the level of presentation and presence of noise, the speech intelligibility seem to decrease with decreasing cut-off frequency in the low-pass filtered conditions and with increasing cut-off frequency in the band-pass filtered conditions (cf. Horwitz et al., 2002; Vickers et al., 2001). Furthermore, when the speech is presented in noise (N), the percent correct score is lower than for the two quiet conditions. As mentioned before, the levels of presentation were chosen to avoid ceiling effects. This seems to have worked out well in the VCV-test, where results are obtained in all conditions for all filter types. In the CUNY's there seems to be certain problems. The QH and the N conditions are working alright. The results from the CUNY-test in the condition N indicate that the signal-to-noise ratio was too hard in the filtered conditions, while it was better in the VCV-test and thus provided results >0% in all filtered conditions. Thus, if a more thorough testing of the signal-to-noise ratio for the CUNY's is made, there should not be any problems to get results that are not hitting the ceiling (or rather the floor) in all of the filtered conditions in the condition N. The QL condition is the toughest to deal with. There is a problem with the overall level in the narrower filtered conditions that results in scores correct that are probably too high in the wider ones. This is the case in both the CUNY-test and the VCV-test. However, it seems difficult to lower the overall level without adding further high frequency emphasis. A larger amount of high frequency emphasis might be a suitable course to solve this complication.

Generally speaking, all of the tests constructed have provided results that are appropriate according to their rationale. The only exception seems to be the QL condition, where a larger high frequency emphasis is proposed to account for the deviations displayed. Furthermore, if a more accurate signal-to-noise ratio is established in the CUNY-test, all present problems seem to be accounted for and solved. What the results from planned testing of both normal hearing and hearing impaired subjects will show, is a different question.

4.6 Acknowledgments

The author would like to thank the following persons, all in someway involved in this thesis, for skillful support and patience: Dr. Teresa Ching, Dr. Harvey Dillon, Scott Brewer, Michael Fisher, Levi Foster, Johannes Lantz and Dr. Eric LePage. He also would like to thank Lone Koch for love and patience.

5. REFERENCES AND LITERATURE

5.1 References

ANSI. (1969). *ANSI S3.5-1969. American National Standard Methods for the calculation of the articulation index.* New York: American National Standards Institute.

ANSI. (1993). *ANSI S3.5 Draft v.3.1.-1993. Proposed American National Standard Methods for the calculation of the speech intelligibility index.* New York: American National Standards Institute.

Arlinger, Stig & Håkan Dryselius. (1990). Speech Recognition in Noise, Temporal and Spectral Resolution in Normal and Impaired Hearing. *Acta Otolaryngologica* Suppl. 469: 30-37.

Byrne, Dennis, Harvey Dillon, Khanh Tran, Stig Arlinger, Kieth Wilbraham, Robyn Cox, Björn Hagerman, Raymond Hetu, Joseph Kei, C. Lui, Jürgen Kiessling, M. Nasser Kotby, Nasser H. A. Nasser, Wafaa A. H. El Kholy, Yasuko Nakanishi, Herbert Oyer, Richard Powell, Dafydd Stephens, Rhys Meredith, Tony Sirimanna, George Tavartkiladze, Gregory I. Frolenkov, Søren Westerman & Carl Ludvigsen. (1994). An international comparison of long-term average speech spectra. *Journal of the Acoustical Society of America* 96 (4): 2108-2120.

Ching, Teresa Y. C. & Harvey Dillon. 1997. Speech Recognition of Hearing-Impaired Listeners: Relations with Frequency Resolution, Temporal Resolution, Context-Utilization Ability and Age. Unpublished, National Acoustic Laboratories.

Ching, Teresa Y. C., Harvey Dillon & Denis Byrne. (1998). Speech Recognition of Hearing-impaired Listeners: Predictions from Audibility and the Limited Role of High-frequency Amplification. *Journal of the Acoustical Society of America* 103 (2): 1128-1140.

Dillon, Harvey. (2001). *Hearing Aids.* Sydney: Boomerang Press.

Hogan, C & Turner C.W. (1998) High-Frequency Amplification: Benefits for Hearing-Impaired Listeners. *Journal of the Acoustical Society of America* 104: 432-441.

Horwitz, Amy R., Judy R. Dubno & Jayne B. Ahlstrom. (2002). Recognition of low-pass-filtered consonants in noise with normal and impaired high-frequency hearing. *Journal of the Acoustical Society of America* 111 (1): 409-416.

ISO. (1998.) *ISO 389-1. Acoustics – Reference zero for the calibration of audiometric equipment – Part 1: Reference equivalent threshold sound pressure levels for pure tones and supra-aural earphones.*

LePage, Eric, Narelle Murray, John Seymour & Dan Zhou. (2001). Transforming hearing conservation into hearing loss prevention: Testing measures of early warning for cochlear hearing loss. *National Acoustic Laboratories Research and Development Annual Report 2000/2001*. Sydney: Australian Hearing.

Moore, Brian C. J. (1995). *Perceptual Consequences of CochlearDamage*. Oxford psychology series 23. New York: Oxford University Press.

Moore, Brian C. J., Martina Huss and Brian Glasberg. (2000). *Diagnosis of Dead Regions*. CD.

Moore, Brian C. J., M. Huss, D. A. Vickers, B. R. Glasberg & J. I. Alcántara. (2000). A Test for the Diagnosis of Dead Regions in the Cochlea. *British Journal of Audiology* 34 (4): 205-224.

Moore, Brian C. J. (2001). Dead Regions in the Cochlea: Diagnosis, Perceptual Consequences, and Implications for the Fitting of Hearing Aids. *Trends in Amplification* 5 (1): 1-34.

Moore, Brian C. J. & José I. Alcántara. (2001). The Use of Psychophysical Tuning Curves to Explore Dead Regions in the Cochlea. *Ear and Hearing* 22 (4): 268-278.

O'Loughlin, B. J. & B. C. J. Moore. (1981). Off-frequency listening: Effects on psychophysical tuning curves obtained in simultaneous and forward masking. *Journal of the Acoustical Society of America* 69 (4): 1119-1125.

Pavlovic, Chaslav V., Gerald A. Studebaker & Robert L. Sherbecoe. (1986). An Articulation Index Based Procedure for Predicting the Speech Recognition Performance of Hearing-impaired individuals. *Journal of the Acoustical Society of America* 80 (1): 50-57.

Pickles, James O. (1995). *An Introduction to the Physiology of Hearing*. London: Academic Press.

Speech and Noise CD for Hearing Aid Evaluation. (2000). Sydney: National Acoustic Laboratories, Australian Hearing.

Vickers, Deborah A., Brian C. J. Moore & Thomas Baer. (2001). Effects of Low-pass Filtering on the Intelligibility of Speech in Quiet for People with and without Dead Regions at High Frequencies. *Journal of the Acoustical Society of America* 110 (2): 1164-1175.

Zwicker, E., & E. Terhardt. 1980. Analytical expressions for critical-band rate and critical bandwidth as a function of frequency. *Journal of the Acoustical Society of America* 68 (5): 1523-1525.

5.2 Literature

Bench, John & John Bamford (Eds). (1979). *Speech Hearing Tests and the Spoken Language of Hearing-Impaired Children*. London and New York: Academic Press.

Bess, Fred H. (1983). Clinical assessment of speech recognition. *Principles of Speech Audiometry*. Ed. By Dan F. Konkle and William F. Rintelmann. Baltimore, Maryland: University Park Press.

Gelfand, Stanley A. (1990). *Hearing. An Introduction to Psychological and Physiological Acoustics*. New York: Marcel Dekker.

Moore, Brian C. J. (1989). *An introduction to the Psychology of Hearing*. London: Academic Press.

Murray, Narelle & Denis Byrne. (1986). Performance of hearing-impaired and normal hearing listeners with various high frequency cut-offs in hearing aids. *Australian Journal of Audiology* 8 (1): 21-28.

Wilber, Laura Ann. (1972). Calibration: Pure Tone, Speech and Noise Signals. *Handbook of Clinical Audiology*. Ed. by Jack Katz. Baltimore, Maryland: Waverly Press.